the plant journal

Glutathione transferases catalyze recycling of auto-toxic cyanogenic glucosides in sorghum

Journal:	*The Plant Journal*
Manuscript ID	TPJ-00899-2017.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	13-Feb-2018
Complete List of Authors:	Bjarnholt, Nanna; Plant Biochemistry Laboratory, Department of Plant and Environmental Sciences, University of Copenhagen; VILLUM Research Center for "Plant Plasticity", Department of Plant and Environmental Sciences, University of Copenhagen Neilson, Elizabeth; Kobenhavns Universitet Det Natur- og Biovidenskabelige Fakultet Crocoll, Christoph; University of Copenhagen, Plant and Environmental Sciences Jørgensen, Kirsten; University of Copenhagen, Department of Plant and Environmental Sciences Motawie, Mohammed; University of Copenhagen, Department of Plant and Environmental Sciences Olsen, Carl; University of Copenhagen, Department of Plant and Environmental Sciences Dixon, David; GSK Edwards, Robert; Newcastle University Møller, Birger; University of Copenhagen, Department of Plant and Environmental Sciences
Key Words:	Cyanogenic glucosides, Sorghum bicolor, Glutathione transferases, Dhurrin, Nitrilases, Ressource allocation

Glutathione transferases catalyze recycling of auto-toxic cyanogenic glucosides in sorghum

Running title: GST-mediated recycling of defense compounds

Nanna Bjarnholt[1,2], Elizabeth H. Neilson[1,2], Christoph Crocoll[2,3], Kirsten Jørgensen[2], Mohammed Saddik Motawia[1,2], Carl Erik Olsen[1,2], David P. Dixon[4,5], Robert Edwards[4,6], Birger Lindberg Møller[1,2].

Affiliations at the time the work was done:

[1]VILLUM Research Center for Plant Plasticity, Department of Plant and Environmental Sciences, University of Copenhagen, Frederiksberg 1871, Denmark

[2]Plant Biochemistry Laboratory, Department of Plant and Environmental Sciences,University of Copenhagen, Frederiksberg 1871, Denmark

[3]DynaMo Center, Department of Plant and Environmental Sciences, University of Copenhagen, Frederiksberg 1871, Denmark

[4]Center for Bioactive Chemistry, Durham University, Durham DH1 3LE, UK

Current affiliations if different from above:

[5]GSK, Stevenage, Hertfordshire SG1 2NY, UK

[6]NU-Agriculture, School of Natural and Environmental Sciences, Newcastle University, Newcastle NE1 7RU, UK.

Contact: Corresponding authors: Nanna Bjarnholt (nnb@plen.ku.dk), Birger Lindberg Møller (blm@plen.ku.dk).

Abstract

Cyanogenic glucosides are nitrogen-containing specialized metabolites that provide chemical defense against herbivores and pathogens via release of toxic hydrogen cyanide. It has been suggested that cyanogenic glucosides are also a store of nitrogen that can be remobilized for general metabolism via a previously unknown pathway. Here we reveal a recycling pathway for the cyanogenic glucoside dhurrin in sorghum (*Sorghum bicolor*) that avoids hydrogen cyanide formation. As demonstrated *in vitro*, the pathway proceeds via spontaneous formation of a dhurrin derived glutathione conjugate, which undergoes reductive cleavage by glutathione transferases of the plant specific lambda class (GSTLs) to produce *p*-hydroxyphenylacetonitrile. This is further metabolized to *p*-hydroxyphenylacetic acid and free ammonia by nitrilases, and then glucosylated to form *p*-glucosyloxyphenylacetic acid. Two of the four GSTLs in sorghum exhibited high stereospecific catalytic activity towards the glutathione conjugate, and form a subclade in a phylogenetic tree of GSTLs in higher plants. The expression of the corresponding two *GSTLs* co-localized with expression of the genes encoding the *p*-hydroxyphenylacetonitrile-metabolizing nitrilases at the cellular level. The elucidation of this pathway places GSTs as key players in a remarkable scheme for metabolic plasticity allowing plants to reverse the resource flow between general and specialized metabolism in actively growing tissue.

Significance statement

Glutathione transferase enzymes (GSTs) are highly abundant in plants, but only very few have been assigned a physiological function. Here we demonstrate that GSTs are key players in a pathway for endogenous recycling of a cyanogenic glucoside defense compound in sorghum, revealing that plants continuously optimize, and can even reverse, their allocation of resources between general and specialized metabolism.

Introduction

Plants produce a range of specialized metabolites to fend off herbivores and pests, to attract pollinators and adapt to a changing environment. One such class of compounds is comprised of the cyanogenic glycosides, which consist of an amino acid derived carbon backbone and a cyanohydrin function stabilized through glucosylation. Cyanogenic glycosides are defense compounds widely distributed in terrestrial plants, including the pteridophytes, gymnosperms and angiosperms (Zagrobelny et al., 2008). In many species these compounds are produced in very high concentrations. Prominent examples are bitter almonds (*Prunus amygdalis*) and seeds of black cherry (*Prunus serotina*), tubers of cassava (*Manihot esculenta*) and the aerial tissues of eucalypts (*Eucalyptus* ssp.), legumes and sorghum (*Sorghum bicolor*). In each case, cyanogenic glycosides comprise several percent of plant dry matter, thereby tying up an appreciable fraction of the plants' carbon and nitrogen pools (Swain and Poulton, 1994; Crush and Caradus, 1995; Gleadow and Woodrow, 2000; Busk and Møller, 2002; Forslund et al., 2004; Jørgensen et al., 2005; Sanchez-Perez et al., 2008). In sorghum, biosynthesis of the cyanogenic glucoside dhurrin (Figure 1) occurs rapidly in young seedlings and in developing grains and in both cases is accompanied by continuous further metabolism of the produced dhurrin. At later developmental stages, the rate of this metabolism exceeds that of biosynthesis, resulting in a net depletion of the total dhurrin pool (Adewusi, 1990; Busk and Møller, 2002; Nielsen et al., 2016). Other cyanogenic species are also known to metabolize cyanogenic glycosides in the course of development. It has therefore often been suggested that in addition to their roles in defense, cyanogenic glycosides serve a dual purpose as storage compounds for reduced nitrogen that can be released for incorporation into general metabolites upon demand (Lieberei et al., 1985; Swain and Poulton, 1994; Jørgensen et al., 2005; Møller, 2010; Neilson et al., 2011; Picmanova et al., 2015; Nielsen et al., 2016). The mechanisms for cyanogenic glycoside recycling and N-remobilization have not been defined to date. However, recently common types of likely derivatives of cyanogenic glycosides have been identified in three unrelated plant species, namely almond, cassava and sorghum. These plant species produce cyanogenic glycosides from four different amino acid precursors, pointing toward the existence of at least one conserved pathway (Picmanova et al., 2015; Blomstedt et al., 2016).

The role of cyanogenic glycosides as defense compounds relies on their bio-activation by specific endogenous β-glycosidases (BGD) to release hydrogen cyanide (HCN), that is toxic to respiring organisms, including plants (Siegien and Bogatek, 2006). HCN is also formed as a by-product of

ethylene biosynthesis in stoichiometric amounts (Peiser et al., 1984). As such, all plants harbor an enzyme system for HCN detoxification and nitrogen recovery. Accordingly, it has been suggested that the recovery of nitrogen from stored cyanogenic glycosides proceeds *via* the bioactivation and detoxification pathways, outlined in Figure 1A, B (Miller and Conn, 1980; Swain and Poulton, 1994). In this pathway, β-cyanoalanine synthase (CAS) catalyzes the incorporation of the released HCN into β-cyanoalanine which is then converted into asparagine and/or aspartate plus ammonia by members of the nitrilase 4 (NIT4) family. In most plants the reaction is catalyzed by a single NIT4 enzyme. However, grasses (Poaceae) possess two NIT4 homologs (NIT4A and NIT4B), with catalytic activity being dependent on heteromer formation between these isoforms (Figure 1B) (Jenrich et al., 2007). In sorghum, three NIT4 isoforms (NIT4A, NIT4B1 and NIT4B2) are present, with the heteromer NIT4A/NITB2 able to hydrolyze *p*-hydroxyphenylacetonitrile, as well as β-cyanoalanine (Figure 1c) (Jenrich et al., 2007). Remarkably, the heteromer NIT4A/NITB1was only found to hydrolyze β-cyanoalanine. Because of the structural resemblance of dhurrin and *p*-hydroxyphenylacetonitrile, it was suggested that the NIT4A/NIT4B2 heterodimer is part of an uncharacterized pathway enabling recycling of nitrogen from dhurrin without releasing toxic HCN within plant cells (Jenrich et al., 2007). In support of this hypothesis, expression of *NIT4A/NIT4B2*, but not of *NIT4B1*, in the developing sorghum grain was correlated with dhurrin depletion. In contrast, the two known β-glucosidases specific for the hydrolysis of dhurrin, dhurrinase 1 and 2, were not expressed (Nielsen et al., 2016).As pointed out by Jenrich et al. (2007), the product of a putative pathway for dhurrin recycling proceeding via the hydrolysis of *p*-hydroxyphenylacetonitrile catalyzed by the NIT4A/NIT4B2 heteromer would be *p*-hydroxyphenylacetic acid (Figure 1c). Previous studies have demonstrated the presence of the corresponding glucoside of this compound, namely *p*-glucosyloxyphenylacetic acid, in sorghum plants as well as in grains (Picmanova et al., 2015; Nielsen et al., 2016). The accumulation of this conjugate was reduced to almost zero in *tcd2* sorghum mutant plants impaired in the last step of the dhurrin biosynthetic pathway (Blomstedt et al., 2016). The evidence for a recycling pathway proceeding *via* non-BGD mediated conversion of dhurrin into *p*-hydroxyphenylacetonitrile, *p*-hydroxyphenylacetic acid and subsequently *p*-glucosyloxyphenylacetic acid is therefore compelling. However, it has not been experimentally verified that the *p*-glucosyloxyphenylacetic acid found in sorghum is indeed derived from dhurrin, and if so, if and how dhurrin can be converted into *p*-hydroxyphenylacetonitrile *in planta*. Here we address these issues and present

evidence of the involvement of a specific class of glutathione transferase enzymes (GSTs) and their substrate, the tripeptide glutathione (GSH), in such a recycling pathway, as outlined in Figure 1c.

Results

Dhurrin accumulation and metabolism in young sorghum plants

To investigate the possible presence of a pathway for conversion of dhurrin to *p*-glucosyloxyphenylacetic acid, the total content of these two metabolites was quantified in young sorghum plants. In a representative experiment the dhurrin content decreased approximately 17 days after sowing, while the content of *p*-glucosyloxyphenylacetic acid increased throughout the experimental period (Figure 2). While the specific number of days that elapsed between sowing and maximal dhurrin content varied between experiments, in all cases the data confirmed that maximal dhurrin content was reached after day 10. Likewise, in all experiments, the replicates sampled around the time of maximal dhurrin content were subject to large variations, as seen in Figure 2, notably between days 13 and 24. Similar large variations in dhurrin content have been determined in developing sorghum grain (Nielsen et al., 2016). This variation in dhurrin content indicates the onset of a change in the relative rates of synthesis vs. metabolism of dhurrin, with the differences between individual plants or grains giving rise to the variability. The total amount of *p*-glucosyloxyphenylacetic acid (abbreviated pGPAAc in figures) began to increase rapidly with plant age around the time where the standard deviations on the measurements of dhurrin content indicated a change in overall metabolism (Figure 2). This is reflected in the embedded table in Figure 2, showing the ratio dhurrin : *p*-glucosyloxyphenylacetic acid. This ratio was constant at the first three data points; at day 10 there was enormous variation and after that the ratio decreased, demonstrating that the rate of *p*-glucosyloxyphenylacetic acid exceeded that of dhurrin in the growth period covered by the last three data points. Neither free *p*-hydroxyphenylacetic acid, nor its suggested precursor, *p*-hydroxyphenylacetonitrile, were identified in the extracts.

To ascertain that *p*-glucosyloxyphenylacetic acid in sorghum is indeed derived from dhurrin, we carried out two different experiments. The first made use of the sorghum mutant line *tcd1* (*total cyanide deficient 1*). Plants of this line are mutated in the first enzyme of the dhurrin biosynthetic pathway (Blomstedt et al., 2012), only accumulating minute amounts of dhurrin and no detectable *p*-glucosyloxyphenylacetic acid (Figure 3A). When a dhurrin solution was administered to excised *tcd1* leaves an appreciable amount of *p*-glucosyloxyphenylacetic acid was accumulated, verifying

their metabolic relationship (Figure 3A, similar data for leaves from older plants are provided in Figure S1). In the second experiment, intact wildtype plants were administered ^{14}C-labeled L-tyrosine at the 2½ leaf stage (5 days old) and the plants harvested 3, 6 and 11 days after uptake. An initial experiment showed that at all time points, the majority of radiolabel was found in the node and in the leaves that were developed at the time of administration. Leaf 2 was fully developed at 5 days and remained green throughout the experiment. From each plant, leaf 2 was therefore harvested, extracted and analyzed separately, and also analyzed in combination with extracts of the remaining plant tissues (respectively "leaf 2" and "total plant" in Figure 3B+C and Figure S2). The total incorporation of label into dhurrin was relatively uniform between replicates (Figure 3C and Figure S2), but decreased over the course of the experiment, demonstrating the continuous metabolism of dhurrin. In the TLC analysis of total plant extracts, the radiolabeled spots pertaining from *p*-glucosyloxyphenylacetic acid were increasing in intensity with time, although at levels too weak to be reliably quantified (Figure S2). In leaf 2 extracts it was possible to quantify the intensity of ^{14}C-labeled *p*-glucosyloxyphenylacetic acid spots, and it was found to increase by approximately 130 % from 3 to 11 days after uptake, while the ratio ^{14}C-dhurrin : ^{14}C-*p*-glucosyloxyphenylacetic acid decreased by a factor 3 (Figure 3C). Remarkably, total ^{14}C in total extracts also decreased over time, demonstrating loss of label to either roots, atmosphere or the insoluble fraction. In the soluble fraction, only very few other compounds than dhurrin and *p*-glucosyloxyphenylacetic acid were found to be labelled (Figure S2), and none of them co-eluted with known compounds.

Identification of GSH conjugation products of dhurrin

When investigating nitrilase activity *in vitro*, a reductant is often added to the assay mixture to stabilize enzyme activity (Jenrich et al., 2007). In previous studies (Jenrich et al., 2007), the classical reductant dithiothreitol (DTT) was found to cause a chemical conversion of dhurrin into *p*-hydroxyphenylacetonitrile. On testing other reductants, we found that β-mercaptoethanol caused the same chemical conversion as DTT, while the tripeptide glutathione (GSH), a ubiquitous antioxidant in plants, undergoes different chemical reactions with dhurrin. As determined by LC-MS/MS analysis of samples containing dhurrin incubated with GSH, the reaction products resulted in two peaks of *m/z* 439 (Figure S3), with this ion producing a major fragment at *m/z* 310 and a minor entity at *m/z* 364. Such neutral losses of respectively 129 or 75 Da are commonly observed in positive mode electrospray when analyzing GS-conjugates by LC-MS/MS (Chen et al., 2008; Xie et al., 2013). The compounds were produced *in vitro*, purified by preparative HPLC and their

structures elucidated by NMR (see the Methods section, Figure S3 and Table S1 for chemical shifts and their assignments). It was determined that GSH had substituted the glucose in dhurrin to form glutathionylated *p*-hydroxyphenylacetonitrile (GS-pOHPACN, Figure 1C). As determined by LC-MS/MS and NMR, the *in vitro* reaction produced a ~50:50 mixture of GS-pOHPACN epimers, indicating that the conjugate formation proceeded by a S_N1 nucleophilic substitution reaction. The two epimers were named GS-pOHPACN 1 and GS-pOHPACN 2, reflecting their order of elution from the HPLC column. No attempts were made to clarify the configuration around the dhurrin-derived chiral carbon. GS-pOHPACN formation was inhibited at pH < 5 and promoted with increasing pH. This is consistent with the previous finding that dhurrin forms a reactive quinone tautomer at high pH, which in the absence of other reactants causes hydrolysis of the *O*-glucosidic bond and the non-enzymatic formation of HCN and *p*-hydroxybenzaldehyde (Mao and Anderson, 1965; Møller et al., 2016). The GS-pOHPACN epimers isomerized rapidly at pH > 5. When one epimer was incubated at pH 7.5 (30 °C), 5 % isomerized into the other within 1 min. After a 30 min incubation, an equilibrium was reached, resulting in a ~50:50 mixture of GS-pOHPACN epimers.

Having identified this GSH-mediated reaction *in vitro*, the sorghum plant extracts (Figure 2) were re-analyzed to determine if these characteristic conjugates could be detected *in vivo*. GS-pOHPACN was present in small amounts in the extracts, but only detectable in all replicates at days 3 (trace), 7 (9 nmol/plant, s.d. 3 nmol/plant) and 24 (190 nmol/plant, s.d. 163 nmol/plant). The measured concentrations were adequate to allow quantification, but not to determine which of the GS-pOHPACN epimers were present *in planta*. Upon revisiting the metabolite analyses data of earlier studies (Nielsen et al., 2016), the conjugate was also detected in small amounts at the end of sorghum grain development, when dhurrin content was rapidly declining (Figure S4).

Dhurrin metabolizing enzyme activities in sorghum seedlings

The discovery of GS-pOHPACN prompted us to investigate whether glutathione transferases (GSTs) were involved in dhurrin recycling. GSTs are known to catalyze a range of conjugation and reductive reactions using GSH as a co-substrate or cofactor (Dixon and Edwards, 2010a; Cummins et al., 2011; Labrou et al., 2015). Soluble protein was extracted from the same batch of sorghum plants as used for the metabolite analysis (Figure 2) and assayed for relevant GST and nitrilase activities. When the plant protein extracts were supplied with GS-pOHPACN 2 as substrate, in the presence of GSH, *p*-hydroxyphenylacetonitrile and *p*-hydroxyphenylacetic acid were produced in the samples (Figure 4B). In the absence of plant extract, no products were detected and no substrate

was consumed. Likewise, incubation with GS-pOHPACN 1 did not give rise to any detectable product formation or substrate consumption (Figure 4A). The incubation time was kept at 5 min to minimize isomerization, and the results therefore demonstrate specific conversion of a single GS-pOHPACN epimer, pointing towards an enzyme catalyzed reaction.

The plant protein extracts were also incubated with dhurrin in the presence of GSH, giving rise to the production of GS-pOHPACN as well as *p*-hydroxyphenylacetonitrile and *p*-hydroxyphenylacetic acid (Figure 4C). The dominant product was *p*-hydroxybenzaldehyde formed from the highly active bioactivation pathway (Figure 4D). This overwhelming dhurrinase activity rapidly consumed the supplemented dhurrin, and it was therefore necessary to use a high concentration of substrate (2.5 mM) and extend the incubation time (45 min) to detect the putative recycling pathway products. This explains why in these samples trace amounts of *p*-hydroxyphenylacetamide were also detected, as this compound is a minor by-product of the nitrilases from sorghum formed *in vitro* (Agerbirk et al., 2008). Under these prolonged assay conditions, in the absence of plant extract, the only product was GS-pOHPACN (Figure 4D), further supporting that conversion of GS-pOHPACN to *p*-hydroxyphenylacetonitrile was not due to a spontaneous chemical reaction. The combined amount of GS-pOHPACN, *p*-hydroxyphenylacetonitrile and *p*-hydroxyphenylacetic acid produced, was much less than the amount of GS-pOHPACN formed in the protein-free control incubations. It was concluded that the high dhurrinase activity present in the extracts prevented the detection of possible enzyme catalyzed formation of GS-pOHPACN from dhurrin. Significantly, the highest levels of GS-pOHPACN determined in the enzyme assays were with protein extracts from those plants where the compound was also found in quantifiable amounts in the metabolite extracts, namely day 7 and day 24 (Figure 4). This indicated that the higher level of GS-pOHPACN found in plants at these two time points was defined by components of the soluble protein fraction.

Novel roles for GSTs in dhurrin recycling

The dhurrin derivatives formed in the plant protein extracts in the presence of GSH revealed the existence of a protein component with the ability to catalyze the reductive cleavage of GS-pOHPACN. Orthologous GSH-dependent activities have previously been determined with members of plant specific subclasses of the GST enzyme super-family (GSTs, EC 2.5.1.18), notably the dehydroascorbate reductases (DHARs) and the lambda GSTs (GSTLs) (Dixon et al., 2002; Dixon and Edwards, 2010b; Lallement et al., 2014b). Thus these enzymes are known to catalyze the

reductive cleavage of GSH conjugates, resulting in removal of GSH from the molecules (see GSTL reaction Figure 1c). In the sorghum genome (http://phytozome.jgi.doe.gov/pz/portal.html#), three sequences for putative *Sb*DHAR enzymes and four putative *Sb*GSTLs have been identified based on *in silico* analysis of GSTs from all published plant genomes (Lallement et al., 2014b). Furthermore, the expression of these genes has been confirmed by transcriptomics in developing sorghum grain (Nielsen et al., 2016). To examine their function, synthetic genes encoding all seven putative GSTs were expressed in *E. coli* and the resulting proteins tested for their ability to metabolize dhurrin and GS-pOHPACN. Only the members of the GSTL class were found to exert activity on these substrates, and as such were selected for further analyses.

The *GSTL*s were cloned from sorghum cDNA to verify that the expressed coding sequences corresponded to those derived from the genomic sequences and transcriptomics experiment. The corresponding enzymes were named *Sb*GSTL1 to *Sb*GSTL4, respectively, and the genes encoding them termed *SbGSTL1* (transcript Sobic.002G421200.2), *SbGSTL2* (transcript Sobic.009G033200.2), *SbGSTL3* (transcript Sobic.001G412700.1) and *SbGSTL4* (transcript Sobic.001G412800.2). Of these, *Sb*GSTL1 and *Sb*GSTL2 were 94 % identical and shared low identity with the other GSTLs (32-51 %) (Figure S4 and Data S1). A phylogenetic analysis of GSTLs from sorghum, wheat, rice, Arabidopsis, poplar and maize identified that *Sb*GSTL1 and *Sb*GSTL2 form a specific subgroup (Figure 5). This suggested that they had evolved into a sub-clade early after the emergence of lambda GSTs. In contrast, *Sb*GSTL3 and *Sb*GSTL4 shared more than 65 % identity with other monocot GSTLs and are apparently members of two different GSTL subgroups. *Sb*GSTL4 had a predicted chloroplast targeting peptide (aa 1-15). Plastid targeted GSTLs have been identified in other plants, e.g. in Arabidopsis (*At*GSTL2 (Dixon et al., 2002)) and poplar (*Pt*GSTL1 (Lallement et al., 2014b)).

The coding sequences of the four *GSTL* genes and that of a truncated version of *SbGSTL4* lacking the chloroplast targeting peptide were expressed in *E. coli* with an N-terminal Strep3-tag and purified. The recombinant enzyme preparations were then tested for their ability to catalyze the GS-pOHPACN cleavage reaction. The experiments were also carried out using dhurrin and the reference compound 4NPG (*S*-(4-nitrophenacyl)glutathione) as substrates. For comparison with the enzymes from sorghum, purified Strep3-tagged GSTLs from *Arabidopsis* (Dixon et al., 2009) and wheat (Dixon and Edwards, 2010b) were also assayed at a fixed, high concentration of GS-pOHPACN (500 µM). The reaction's pH dependency was determined for *Sb*GSTL2 using GS-

pOHPACN 2 as substrate. While the reaction rate was similar at pH 6 - 6.5, it was found to increase between pH 6.5 - 8. Most plant GSTs, including all lambda GSTs investigated to date, have been found to be soluble enzymes. Furthermore, in localization experiments the majority of GSTs have been found to be cytosolic, with a few exceptions mainly ascribed to predicted target peptides on the enzymes (Dixon et al., 2002; Dixon et al., 2009; Lan et al., 2009; Dixon and Edwards, 2010b; Liu et al., 2013; Lallement et al., 2014b). Here, only one enzyme had a predicted target peptide (*Sb*GSTL4), and therefore the remaining experiments were carried out at pH 7.5 to ensure near optimal conditions for the enzyme activity while remaining at a realistic pH for a cytosolic reaction. The GSTL catalyzed reductive cleavage reaction rate depends on the binding of the co-factor GSH, along with the conjugate molecule. To monitor specifically the affinity of the GSTL enzymes for the conjugate molecule, the assays were performed in the presence of a large excess of GSH. To minimize the effects of GS-pOHPACN isomerization, incubation times were minimized and the isomerization reaction excluded from the calculations of the kinetic parameters. The resulting apparent K_m and k_{cat} values with respect to the two different GS-pOHPACN epimers are presented in Table 1.

Of the eleven different GSTL-enzymes tested for their ability to convert GS-pOHPACN to *p*-hydroxyphenylacetonitrile, nine were found to be active (Table 1 and Table S2). *Sb*GSTL3 and *Ta*GSTL1 were not able to metabolize GS-pOHPACN. While the three catalytically active *Sb*GSTLs had K_m values in the same range, the catalytic efficiencies (k_{cat}/K_m) for *Sb*GSTL1 and *Sb*GSTL2 were 10-100 fold higher than for *Sb*GSTL4 and Δ*Sb*GSTL4. The turnover numbers determined for wheat and *Arabidopsis* GSTLs at high substrate concentrations were also low compared to k_{cat} values for *Sb*GSTL1 and *Sb*GSTL2. Importantly, *Sb*GSTL1 and *Sb*GSTL2 were the only two enzymes demonstrating clear preference for a specific GS-pOHPACN epimer, namely GS-pOHPACN 2, as also determined in the sorghum plant protein extracts (Figure 4A and B). The compound 4-NPG is structurally similar to GS-pOHPACN and was previously identified as a substrate for the human GST, omega GST 1-1, which also carries out reductive cleavage reactions (Board et al., 2008). In the present study, the activity towards 4-NPG was used to verify that the recombinant *Sb*GSTL3 was catalytically active and to further substantiate that GS-pOHPACN is likely to be a physiological substrate of *Sb*GSTL1 and *Sb*GSTL2. Using 4-NPG as substrate for *Sb*GSTL1 and *Sb*GSTL2, V_{max} was not reached even at a substrate concentration of 500 μM, indicating that the K_m for 4-NPG was much higher than for both of the GS-pOHPACN epimers. The turnover numbers were respectively 466 min^{-1} (s.d. 13 min^{-1}) and 402 min^{-1} (s.d. 94 min^{-1}) at

this substrate concentration. *Sb*GSTL3 and *Sb*ΔGSTL4 both displayed relatively high K_m (33 μM (s.d. 18 μM) and 54 μM (s.d. 31 μM), respectively) and relatively low catalytic efficiencies (both 2 min^{-1} μM^{-1} (s.d. 1 min^{-1} μM^{-1})) with 4-NPG as substrate.

In the bacterium *Sphingomonas chlorophenolica*, GSTs catalyze the reductive cleavage of pentachlorophenol, in a two-step dechlorination reaction. In the first step, a chlorine atom is replaced by GSH, which is then subsequently removed in a reductive cleavage reaction (McCarthy et al., 1996; Huang et al., 2008). Consequently, these GSTs are both glutathionylating and de-glutathionylating enzymes. Dhurrin was therefore tested as a substrate for the *Sb*GSTLs in the presence of GSH to determine whether GS-pOHPACN is an intermediate in a single enzyme catalyzed conversion of dhurrin to *p*-hydroxyphenylacetonitrile. When the *Sb*GSTLs were supplied with dhurrin, no *p*-hydroxyphenylacetonitrile-forming activity could be detected during the short term incubations giving rise to the data in Table 1, demonstrating that dhurrin was not a substrate. Formation of *p*-hydroxyphenylacetonitrile was detected following prolonged incubation due to the chemical conversion of dhurrin to GS-pOHPACN. The *Sb*GSTLs were also assayed in combination with each other to test the possibility that heteromer formation is essential for their catalytic activity, as observed for the sorghum NIT4 nitrilases (Jenrich et al., 2007). When the enzymes were tested in combination, the total *p*-hydroxyphenylacetonitrile forming activity was equal to the sum of the activities of the individual enzymes.

Site of expression of the genes involved in the recycling pathway

In situ-PCR experiments (Koltai and Bird, 2000; Jørgensen et al., 2005; Jørgensen et al., 2011) were carried out to investigate in which cell types the sorghum *GSTL*s and *NIT4*s were expressed. Experiments were performed using sorghum plants grown for 7 and 10 days. At these time points, enzymes involved in dhurrin recycling would be expected to be active and increasing in amount (Figure 2 and Figure 4), with the corresponding transcripts therefore likely to be present. Figure 6 presents images from two different representative experiments (supported by additional images in Figures S6 and S7). Images 1.A-1.D display when and where the GS-pOHPACN metabolizing *GSTL*s were expressed, and images 2.A-2.D visualize *NIT4* gene expression compared to *SbGSTL1*. Expression of mRNAs encoding *SbGSTL1* or *SbGSTL2* in the third leaf from 10-day-old plants were always found exclusively in the sheath bundle cells and in parenchyma cells surrounding the vascular bundle (Figure 6, 1A and 1B). In contrast, *SbGSTL4* expression as observed in the leaf of a 7-day-old plant was more dispersed, with the respective mRNA found in the epidermis and cortex

as well as scattered in parenchyma cells around the vascular bundles (Figure 6, 1C). Importantly, the *SbGSTL4* transcript was not detectable at day 10 whereas *SbGSTL1* was detected both at day 7 and 10, in the same cells (Figure S6). The nitrilases *SbNIT4A* (Figure 6, 2B) and *SbNIT4B2* (Figure 6, 2D) were consistently found to be expressed in the same cells as *SbGSTL1* (Figure 6, 2A). In contrast, mRNA from the nitrilase *SbNIT4B1*, which is not involved in conversion of *p*-hydroxyphenylacetonitrile to *p*-hydroxyphenylacetic acid (Figure 6, 2C), could not be detected.

Discussion

In this study we have elucidated an endogenous pathway for recycling of the cyanogenic glucoside dhurrin in sorghum. The pathway proceeds via the intermediates GS-pOHPACN, *p*-hydroxyphenylacetonitrile and *p*-hydroxyphenylacetic with concomitant release of ammonium ions (Jenrich et al., 2007) and presumably glucose. The identification of the minimal set of genes and enzymes enabling recovery of nitrogen from dhurrin in sorghum finally corroborates the hypothesis that the defense compounds cyanogenic glycosides serve additional roles as storage compounds for reduced nitrogen.

Endogenous metabolism of dhurrin in sorghum was previously observed (Adewusi, 1990; Busk and Møller, 2002; Nielsen et al., 2016) and was proposed to involve the action of the *Sb*NIT4A/NIT4B2 heteromer, that catalyzed conversion of *p*-hydroxyphenylacetonitrile into *p*-hydroxyphenylacetic acid (Jenrich et al., 2007). In the GSH-dependent pathway identified here, the conversion of dhurrin to the *Sb*NIT4A/NIT4B2 substrate is predicated on dhurrin first reacting non-enzymatically with GSH to form a conjugate that then undergoes reductive cleavage catalyzed by a GSTL. In the plant cytosol, the GSH concentration has been reported to be 0.2-3 mM (Meyer et al., 2001; Hartmann et al., 2003), making it readily available for participation in metabolic processes. Although GSH is normally thought to be involved in the detoxification of xenobiotics and in maintaining redox balance, it is also known to be a building block in biosynthesis of specialized metabolites (Geu-Flores et al., 2009; Geu-Flores et al., 2011; Kobayashi et al., 2011; Su et al., 2011; Fedrizzi et al., 2012). Molecules conjugated with GSH are rarely detected in plants, but have for example been shown to be intermediates in the biosynthesis of glucosinolates and camalexin. The involved GS-conjugates were only accumulated and detected when the biosynthetic pathways were truncated (Geu-Flores et al., 2009; Geu-Flores et al., 2011), confirming that their formation is transient. In the current study, small amounts of GS-pOHPACN were only detected at very specific time points

although dhurrin metabolism took place at all time points studied (Figure 2, Figure 4). A similar situation was observed in the developing sorghum grain (Figure S4) (Nielsen et al., 2016).

The involvement of GSTs in dhurrin metabolism adds new insight into the multiple functions of these enzymes *in planta*. GSTs are diverse and highly abundant in cereals, typically encoded by between 50-80 genes in higher plants (Dixon and Edwards, 2010a; Jain et al., 2010; Chi et al., 2011; Rezaei et al., 2013). Plant GSTs are divided into 14-15 classes, of which six are plant specific, a few are only present in early plant lineages, and one is an evolutionarily distinct microsomal class of membrane bound enzymes conserved across kingdoms (Jain et al., 2010; Cummins et al., 2011; Liu et al., 2013; Lallement et al., 2014a). Despite the abundance and diversity of plant GSTs and many reports about their *in vitro* activities and role in herbicide detoxification, knowledge about their *in vivo* functions and physiological roles is scarce (Dixon et al., 2010; Cummins et al., 2011; Labrou et al., 2015). Classical GST enzymes catalyze a glutathionylation reaction proceeding *via* the nucleophilic attack of the thiol group of glutathione (GSH). Across all kingdoms GST classes exist where the classical catalytic residue (typically serine in plants) has been replaced by a cysteine (Cys-GSTs), generally rendering the enzymes unable to catalyze glutathionylation reactions (Dixon et al., 2002; Board et al., 2008; Xun et al., 2010; Lallement et al., 2014a; Lallement et al., 2014b). In plants, DHARs and GSTLs are such Cys-GSTs. While they do not catalyze the conjugation of xenobiotics with GSH they do have roles in redox homeostasis. For example the DHARs function in the recycling of ascorbate (DHAR = DeHydroAscorbate Reductase) (Urano et al., 2000; Dixon et al., 2002) and/or GSH (Rahantaniaina et al., 2017). In contrast, no specific *in vivo* function has previously been assigned to GSTLs. Likewise, cleavage of GS-conjugates has not previously been demonstrated as a GST catalyzed reaction operating *in planta*. It was therefore surprising to discover sorghum GSTLs committed to deglutathionylation of GS-pOHPACN to *p*-hydroxyphenylacetonitrile in endogenous plant metabolism. GSTLs from wheat, *Arabidopsis* and poplar as well as other Cys-GSTs from mammals and microbes catalyze the reductive cleavage of aromatic GSH conjugates (McCarthy et al., 1996; Tocheva et al., 2006; Board et al., 2008; Dixon and Edwards, 2010b; Meux et al., 2011; Lallement et al., 2014b). Similarly, the majority of tested GSTLs in this study catalyzed the cleavage of GS-pOHPACN to *p*-hydroxyphenylacetonitrile (Table 1). Uniquely, *Sb*GSTL1 and *Sb*GSTL2 discriminated between GS-pOHPACN 1 and 2, implying that they have evolved to specifically catalyze metabolism of one of the two epimers. Furthermore, their catalytic efficiencies for the preferred substrate were relatively high compared with the other GSTLs tested. In the study of *in vitro* activity of the three GSTLs found in poplar, the

highest catalytic efficiency was found towards the physiologically relevant substrate GS-quercetin (Lallement et al., 2014b). With the most efficient *Pt*GSTL reaching k_{cat}/K_m of 10 $\mu M^{-1}min^{-1}$ for this substrate, *Sb*GSTL1 and *Sb*GSTL2 are significantly more efficient at converting GS-pOHPACN. The position of *Sb*GSTL1 and *Sb*GSTL2 in the phylogenetic tree of GSTL sequences (Figure 5) supports that these two enzymes have evolved to perform a function not relevant in the other plant species represented in the tree. Dhurrin is present in wheat (Winterberg et al., 2014), but in much reduced amounts compared to sorghum, whereas the remaining plants species do not produce cyanogenic glucosides.

The *in situ*-PCR experiments with *Sb*GSTL and *Sb*NIT encoding genes demonstrated similar expression patterns of the mRNAs encoding the enzymes involved in the recycling pathway leading to production of *p*-hydroxyphenylacetic acid (*Sb*GSTL1, *Sb*GSTL2, *Sb*NIT4A, *Sb*NIT4B2) (Figure 6). In contrast, mRNAs encoding the non-related *Sb*GSTL4 and *Sb*NIT4B1 were either localized to different cells than those encoding the genes of the pathway, or they were not detected at the selected time points. In transcriptome analysis of the developing sorghum grain, we found that expression of *SbNIT4A* and *SbNIT4B2* (Nielsen et al., 2016) and accumulation of GS-pOHPACN (Figure S4) increased with dhurrin depletion. Of the 59 GST genes expressed in the developing sorghum grain, expression of *SbGSTL1* showed the highest correlation with the expression profiles of *SbNIT4A* and *SbNIT4B2*. The localization of the dhurrin metabolizing enzymes in and around cells of the vascular bundle is consistent with the endogenous recycling pathway providing ammonia and possibly glucose, for distribution within the plant for use in general metabolism. Images of *SbGSTL1* expression were generally easier to obtain than those of *SbGSTL2* expression, indicating that the latter was more weakly expressed, or not always expressed as early as *SbGSTL1*. The sorghum protein activity assays showed that GS-pOHPACN reductive cleaving activity is present from an early stage in seedling development (Figure 4). Presumably, this early activity can be attributed to the clearly expressed *Sb*GSTL1 with expression of the more efficient *Sb*GSTL2 being turned on or increased around the time of maximal dhurrin accumulation where the rate of dhurrin recycling exceeds that of biosynthesis. In the developing sorghum grain, transcripts of both *SbGSTL1* and *SbGSTL2* were relatively highly abundant. However, *SbGSTL2* expression was low in the first phase of grain development and increased dramatically towards grain maturation (Nielsen et al., 2016). Possibly, *Sb*GSTL1 has initially evolved to detoxify and recycle misplaced dhurrin, with the more efficient isoform eventually evolved to be the main driver of purposeful dhurrin recycling.

Neither *Sb*GSTL3 nor *Ta*GSTL1 were able to catalyze the cleavage of GS-pOHPACN, although they were found to act on other aromatic GS-conjugates (Table 1) (Dixon and Edwards, 2010b). Interestingly, these two enzymes share 77-90 % identity with *Os*GSTL1 and *Zm* In2-1 (Figure S4). Together these four enzymes form a distinct cluster in the phylogenetic tree (Figure 5). This suggests that they share a function which is conserved across monocots. A striking feature of these four enzymes is that they contain an additional cysteine three residues upstream from the catalytic cysteine (Figure S4). When this additional cysteine residue was mutated in *Ta*GSTL1, the activity towards both aromatic and non-aromatic substrates increased (Dixon and Edwards, 2010b). Its presence is therefore likely to be responsible for the complete lack of activity towards GS-pOHPACN. *Pt*GSTL2 and *Pt*GSTL3 also have an additional cysteine at this position, but are phylogenetically more closely related to other dicot GSTLs, from Arabidopsis.

The epimer specificity of *Sb*GSTL1 and *Sb*GSTL2 towards GS-pOHPACN could suggest that *in planta*, the preceding conversion of dhurrin to GS-pOHPACN is also catalyzed by an enzyme that would usually lead to the formation of only one epimer. The sorghum GSTLs did not catalyze such a reaction, when tested individually or in combination. However, the sorghum grain transcriptome analyses identified four GST transcripts as being highly abundant and strongly correlated with the expression profiles of *SbNIT4A/SbNIT4B2* (Nielsen et al., 2016). These will be subject of future studies to test their ability to produce GS-pOHPACN from dhurrin. It is also possible that the conjugation of dhurrin to GS-pOHPACN is a strictly chemical process. Although this reaction produces a 50:50 mixture of (*R*)- and (*S*)-GS-pOHPACN, the epimers were found to isomerize relatively quickly. If the production of GS-pOHPACN is non-enzymatic, the epimer which is not favored as a substrate by the GSTLs might be expected to accumulate in preference to the other epimer. To date, we have been unable to determine the precise stereospecific configuration of the minute amounts of GS-pOHPACN detected in the plant extracts. Alternatively, GS-pOHPACN may be produced by the conjugation of a dhurrin derivative. Previously, three dhurrin-glucosides and two presumed caffeic acid esters of dhurrin have been identified in sorghum (Selmar et al., 1996; Picmanova et al., 2015), and one of these may serve as substrate for a conjugating GST.

The identified recycling pathway leads to production of *p*-hydroxyphenylacetic acid, which is subsequently glucosylated. This was demonstrated by the production of *p*-glucosyloxyphenylacetic acid from dhurrin administered to leaves of the *tcd1* sorghum mutant (Figure 3A), and from radiolabeled dhurrin formed upon administration of ^{14}C-labelled tyrosine to wildtype sorghum

plants (Figure 3B+C). Not all recycled dhurrin is converted to this glucosylated product (Figure 2). Other presumed products of recycling pathways for cyanogenic glycosides have been identified in sorghum, cassava and almond (Picmanova et al., 2015; Blomstedt et al., 2016). The ^{14}C-experiment revealed only minor production of other labelled compounds than dhurrin and *p*-glucosyloxyphenylacetic acid (Figure S2), none of which could be identified. However, this experiment was terminated at 11 days, and it is therefore possible that these other presumed products of dhurrin metabolism are produced later. The total amount of ^{14}C in the plant extracts decreased by 40 % over the course of 8 days (Figure 3C). This demonstrates that not only the nitrogen, but also the carbon originally invested in dhurrin is recycled, e.g. into non-methanol extractable polymers such as protein and lignin, or into root biomass. This in turn indicates that the remaining fraction of *p*-glucosyloxyphenylacetic acid serves a specific purpose. *p*-Hydroxyphenylacetic acid has been shown to have growth promoting and other auxin-like effects in plants and algae (Abe et al., 1974; Fries and Iwasaki, 1976; Ding et al., 2008; Simon and Petrášek, 2011), and auxin homeostasis is controlled by glucosylation (Korasick et al., 2013). Consequently, it is plausible that *p*-glucosyloxyphenylacetic acid is an inactive storage form of a signaling compound in sorghum.

Several cyanogenic plant species are polymorphic for the production of cyanogenic glucosides or exist as naturally mutated sub-species that do not accumulate them (Crawford-Sidebotham, 1972; Gleadow et al., 2003; Sanchez-Perez et al., 2008). These defense compounds have therefore traditionally not been considered as compounds contributing to other aspects of plant fitness and physiology, but this assumption has been questioned (Swain and Poulton, 1994; Neilson et al., 2013; Gleadow and Møller, 2014). The use of chemically induced mutations or biotechnological approaches to knock out or reduce cyanogenic glucoside biosynthesis in sorghum and cassava has demonstrated that under specific developmental conditions these plants may show impaired growth (Jørgensen et al., 2005; Blomstedt et al., 2012). The identification of a dhurrin recycling system for recovery of compounds useful for general metabolism, as well as production of a metabolite with putative roles in signaling corroborates the proposed multi-functionality of cyanogenic glucosides and exemplifies how general and specialized metabolism in plants should be envisioned as one highly dynamic and integrated metabolic system. The catalytic repertoire of the GST super family of enzymes is extremely versatile (Dixon and Edwards, 2010a; Dixon et al., 2010; Labrou et al., 2015) and the number of identified plant GST functions outside of detoxification and redox homeostasis is increasing. The identification of GST involvement in cyanogenic glucoside recycling

underpins that members of this enzyme family should be considered as potential key players in the intimate interplay between general and specialized metabolism.

Experimental procedures

Plant material and sampling

The *S. bicolor* plant material used for all experiments except administration of unlabeled dhurrin was the cultivar SS100. Seeds were soaked in water for 24 h and sown on top of wet gauze-covered soil (to ensure uniform length of the mesocotyls). The pots were covered with perforated cling film and germinated in a climate chamber (28 °/25 °C d/n, 80 % humidity, light intensity 480 µmol m^{-2} s^{-2}). After 3 d the cling film and non-germinated seeds were removed. Furthermore, preliminary results showed that at the early stages, dhurrin content was positively correlated with plant size. Therefore the population was trimmed by removing both the most developed and least developed plants and plants that differed visually from the majority (curled leaves, torn leaves etc.). Each sample replicate consisted of all above ground material from a single plant, *i.e.* hypocotyl, node, coleoptile and all leaves, leaf sheaths and possible crown shoots. For administration of dhurrin, seed of the EMS mutant total cyanide deficient (*tcd1*) line, and as control the parent line (Blomstedt et al., 2012), were soaked in water for 24 h, sown on wet soil and covered by a thin layer of moist soil and grown in a greenhouse (28°/25°C d/n). In both experiments, the harvested plants were visually selected to be at an average growth stage of all plants in the pot, *i.e.* an average number of fully unfolded leaves and average height. All samples were immediately frozen in liquid nitrogen at harvest.

Metabolism of dhurrin and L-[U-^{14}C]-tyrosine in sorghum leaves

Dhurrin was administered to leaves of the sorghum *tcd1* mutant line and the parent line. The third and second leaves of 7 and 14 day old plants were excised a few mm below the leaf blade and immediately placed in tubes containing either 500 nmol dhurrin in 50 µl water or 50 µl water, with the few mm of leaf sheath carefully immersed in the solution, and placed under ventilation in a fume hood to promote uptake. As t = 0 controls, leaves of each plant line/age/leaf number were frozen in liquid nitrogen at harvest. All experiments were carried out in triplicate. After uptake, the leaves were transferred to larger tubes containing 500 µl water. The lids were closed, and the leaves allowed to metabolize for 5 h, before they were removed from the water, frozen in liquid nitrogen and subsequently extracted and analyzed by LC-MS as described below. For administration of L-[U-^{14}C]-tyrosine, SS1000 seedlings were removed from the soil 5 d after sowing (2½ leaf stage) and the roots cut approximately 2 cm below the seed to facilitate reproducible uptake across replicates.

Each plant was placed in a tube with the cut roots immersed in 50 µl water containing 2.5 µCi of L-[U-^{14}C]-tyrosine (391 mCi/mmol, Larodan Fine Chemicals). Control plant roots were immersed in water containing an identical concentration of unlabeled L-tyrosine. After uptake, 50 µl water was added to the tubes for rinsing, and when this was also taken up the plants were transplanted back into wet soil to resume growth. The plants were initially affected by the procedure, but after a few days resumed growth and grew normally for the remainder of the period, only 1-2 days delayed compared to undisturbed plants. The plants were harvested in triplicate on day 3, 6 and 11 after uptake and separated into leaf 2 and the remaining tissue, and each part extracted separately, as described below. For determination of total ^{14}C and labeled dhurrin, aliquots of each plant part were combined. Total-^{14}C was measured on a Trilux 1450 MicroBeta liquid scintillation counter (Wallac). Total dhurrin was visualized by TLC-analysis using Silica Gel 60 F_{254} plates from *Merck* and ethylacetate:acetone:water:formic acid (20:70:20:0.5) as elution solvent. All individual extracts were also analyzed by this TLC-system and in one allowing separation of *p*-hydroxyphenylacetonitrile, *p*-hydroxyphenylacetic acid and *p*-hydroxybenzaldehyde using a mobile phase of toluene:ethylacetate:methanol:formic acid (83:17:2:0.5). After sample loading and drying, all plates were initially eluted 2 cm in 100 % MeOH to focus the origin, and dried again before elution in separation solvent. Radiolabeled products were visualized by exposure of phosphor imager screens from Molecular Dynamics which were developed using a STORM 840 phosphor imager (Amersham Biosciences). Spots were quantified using the Imagequant software (Amersham Biosciences).

Extraction of metabolites

The frozen plant samples were extracted 2x3 min in boiling 85 % MeOH containing 0.5 % formic acid to stabilize dhurrin and prevent the GSH-dhurrin chemical reaction. The extraction solvent was added straight to the frozen samples which were immediately transferred to a boiling water bath; nodes and leaf sheaths from large plants were first coarsely ground in liquid nitrogen to ensure complete extraction and immediate quenching of metabolism. The extracts were decanted off and kept at -20° C until analysis.

Chemical standards and substrates

p-Hydroxyphenylacetonitrile, *p*-hydroxyphenylacetic acid, *p*-hydroxyphenylacetamide and *p*-hydroxybenzaldehyde were purchased from Sigma, and dhurrin (Møller et al., 2016) and 4-NPG

(Board et al., 2008) synthesized as described. GS-pOHPACN was synthesized by mixing approximately 80 mmol dhurrin 1:1 with reduced GSH (Sigma) in 200 mM KP_i (adjusted to pH 8 with NaOH) and shaken for 3 h at 37 °C. The resulting mixture was pre-fractioned on Strata-X 33 µm Polymeric reversed phase SPE columns (Phenomenex): 3 ml columns containing 500 mg column material were conditioned with 10 ml methanol and equilibrated with 10 ml mQ-water, 7.5 ml crude GS-pOHPACN mix was applied to each column, followed by 10 ml mQ-water for washing, and the GS-pOHPACN was eluted in 5 ml methanol. This eluate was fractionated on a preparative HPLC from Shimadzu consisting of LC-20AT quarternary pump, CBM-20 communication module, SIL-10AP autosampler, SPD-M20A diode array detector, FRC-10 A fraction collector and two sample coolers to keep samples and fractions cold. The column was a Zorbax SB-C18, 9.4 mm i.d, 150 mm, 5 µm particles (Agilent). The two mobile phases were: A = 0.1 % formic acid, and B = 80 % acetonitrile, 0.1 % formic acid, the flow rate was 7 ml/min, and the gradient program was as follows: 0-10 min: B =10 %, 10-12.5 min: B = linear gradient 10-100 %, 12.5-15 min: B = linear gradient 100-10 %, followed by 10 min equilibration. The yield was on the order of a few percent, but the procedure was not optimized and could most likely benefit from a better pH control during incubation. The GS-pOHPACN structures were verified by LC-MS (Figure S3) and NMR. The ^{1}H and ^{13}C NMR spectra were recorded on a Bruker Advance 400 spectrometer at 400 and 101 MHz, respectively. NMR peak assignments were established by NOESY, COSY, HSQC and HMBC methods. For full assignments of the chemical shifts of the two GS-pOHPACN epimers, see Table S1. In brief, the chemical shifts for the GSH part of the molecules were consistent with those previously reported (Geu-Flores et al., 2009). Furthermore, the ^{1}H NMR data of the dhurrin derived part of the molecules revealed that the aromatic protons appeared as a characteristic AB-system of multiplets at δ 7.41 & δ 6.94 ppm (d, J = 8.6 Hz) for epimer 1 and at δ 7.40 & δ 6.94 ppm (d, J = 8.6 Hz) for epimer 2. This demonstrates the maintenance of the aromatic substitution pattern of the aglycones of (*S*)-dhurrin and its epimer (*R*)-taxiphyllin (Seigler et al., 2005; Møller et al., 2016). On the other hand, the epimeric proton at the CN-bearing carbon (Figure S3) for both epimers was more shielded and appeared as a singlet resonating upfield (δ 5.25 and δ 5.24) compared to dhurrin and taxiphyllin (δ 5.91 and δ 5.79), demonstrating that this carbon is linked to the S-atom of the glutathione molecule. That the CN group is retained in the structure was documented by its diagnostic chemical shift at δ 120.1 ppm for both epimers.

Extraction of soluble protein from sorghum

For each time point, several plants were pooled, ground in liquid nitrogen and added 10 % (m/m) polyvinyl poly pyrrolidone (PVPP from Sigma), followed by approximately 3 volumes (m/v) ice cold 50 mM KP_i (adjusted to pH 7.5 with NaOH) containing 3 mM ethylenediamine-tetra-acetic acid (EDTA), 1 mM phenylmethylsulfonyl fluoride (PMSF) and 1 mM GSH to keep nitrilases reduced. The slurries were filtered through muselin gauze and the insoluble fractions removed by centrifugation, first 20 min at 5,000 g, followed by 30 min at 100,000 g of the supernatants from the first centrifugation. The final supernatants were desalted with PD10 columns (GE healthcare), following the protocol using 5 mM KP_i (adjusted to pH 7.5 with NaOH) with 1 mM GSH as desalting buffer. The extracts were either used immediately or frozen in liquid nitrogen and stored at -80 °C until assayed.

Sequence verification and analyses

The previously identified sequences for three putative *Sb*DHAR enzymes and four putative *Sb*GSTLs (see accession numbers below) were tentatively confirmed by search in EST databases, and the expressed GSTL sequences were verified by cloning from cDNA (3 day old seedlings, iScript cDNA synthesis kit from BioRad, primer pair sequences in Table S3). The online software tool TargetP 1.1 (http://www.cbs.dtu.dk/services/TargetP/) (Nielsen et al., 1997; Emanuelsson et al., 2000) identified a chloroplast targeting signal (first 15 aa) in the translated protein sequence of MF770506 (*Sb*GSTL4) with high probability (cTP = 0.961, RC = 1). Phylogenetic analysis was performed with the software MEGA 6.06 (Tamura et al., 2013): GSTL protein sequences were aligned using MUSCLE (Edgar, 2004) at default settings (Data S1), and the alignment was used to calculate overall sequence percent identities between proteins using the software CLC Main Workbench 7.02 (Figure S5). The same full length sequences were aligned with *At*DHAR1 as outgroup for construction of a maximum likelihood phylogenetic tree by MEGA 6.06 (Data S2). The Analyses Preferences function was first used to decide which model is considered to describe the observed substitution pattern best (the model with the lowest BIC score (Bayesian information criterion)), and the tree was then constructed using 1000 bootstraps, JTT model and a discrete gamma distribution.

Expression and purification of enzymes

Synthetic genes, codon optimized for expression in *Escherichia coli* (*E. coli*), of the three *SbGSTDHAR* sequences, four *SbGSTL* sequences and a truncated version of *SbGSTL4*, *SbΔGSTL4*

with the first 15 aa removed, were purchased from GenScript. These were subcloned into the *E. coli* pET-STRP3 expression vector to add a Strep3 tag to the proteins (Brazier-Hicks et al., 2007), expressed in BL21(DE3) cells (Novagen) and purified as described (Dixon et al., 2009) with the only modifications that the cells were grown in TB broth, and the pellets resuspended in 50 mM KP_i pH (adjusted to 7.5 with NaOH).

Enzyme assays and determination of Michaelis-Menten parameters

The protein content of soluble plant protein samples was estimated from the absorbance at 280 nm measured with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies) and the concentrations adjusted to 1 mg/l in desalting buffer. Of this, 80 µl was mixed with 10 µl substrate solutions (25 mM dhurrin or 2 mM GS-pOHPACN 1 or 2) and 10 µl 30 mM GSH in 40 mM KP_i pH (adjusted to 7.5 with NaOH) to make the final concentrations respectively 2.5 or 0.2 mM (substrate), 3 mM (GSH) and 8 mM (KP_i). The samples were incubated at 37° C for 5 or 45 min and the assays stopped by adding 1 volume MeOH containing 1 % formic acid. The samples were analyzed by LC-MS. GSTL enzyme activity assays were carried out in a modified version of the coupled assay previously described (Edwards and Dixon, 2005), using a Shimadzu UV-2550 UV-VIS spectrophotometer. The activities of the enzymes were measured by the consumption of NADPH (340 nm, $\varepsilon = 6200\ M^{-1}\ cm^{-1}$) by glutathione reductase as it reduces the oxidized GSH produced when the GSTLs cleave GSH conjugates, using the following solutions: A. 200 mM KP_i + 4 mM EDTA (adjusted to pH 7.5 with NaOH), B. 2.5 mM NADPH in 0.1 % $NAHCO_3$, C. 20 mM GSH (adjusted to pH 6 with NaOH), D. 6 u/ml GSH reductase (EC 1.6.4.2 from Sigma) in solution A. In a plastic cuvette 600 µl A (30 °C) and 100 µl each of B, C and D was mixed and allowed to incubate at 30 °C to reduce any oxidized GSH present. After 2 min, 50 µl enzyme solution was added and incubated for an extra 20 sec, after which 50 µl substrate solution was added, and the change in absorbance was measured for 2 min. In the assays, the amount of enzyme was between 0.2-15 µg and substrate concentrations were between 1-250 µM. For the reference compound 4NPG the general procedure was the same, but the reaction rate was measured as the change in 4NPG-concentration at 305 nm ($\varepsilon = -1100\ M^{-1}\ cm^{-1}$) (Board et al., 2008), the mixture was 800 µl A, 100 µl B, 50 µl enzyme, 50 µl substrate, and the substrate concentrations were 20-500 µM. The absorbance at 280 nm of dilution series of enzyme were measured using the NanoDrop spectrophotometer, and extinction coefficients of the Strep3 tagged proteins were estimated using the ProtParam tool found at http://web.expasy.org. The apparent k_{cat} and K_m values were calculated with the SigmaPlot 12.5

software using a non-linear single substrate Michaelis-Menten regression model. The activity assays of *Sb*GSTL1-*Sb*GSTL4 towards GS-pOHPACN and 4-NPG were performed in respectively three and two series of technical replicates.

LC-MS analyses of plant extracts and enzyme assay samples

Plant and enzyme assay samples were analyzed by two different LC-MS procedures. All samples were diluted appropriately in mQ-water to make the concentrations fall within the standard curves and filtered through 0.45 µm Multiscreen HTS HV filters (Millipore). For detection of dhurrin, *p*-glucosyloxyphenylacetic acid and GS-pOHPACN in plant extracts and for analysis of the purified GS-pOHPACN epimers, samples were analyzed as previously described (Picmanova et al., 2015) using an Agilent 1100 Series LC (Agilent Technologies) coupled to a Bruker HCT-Ultra ion trap mass spectrometer (Bruker Daltonics) in positive mode ESI (ElectroSpray Ionization). The compounds were quantified using standard curves in the range 0.1-400 µM. To study dhurrin metabolism in solutions of plant protein extracts (Figure 4), dhurrin, GS-pOHPACN, *p*-hydroxyphenylacetonitrile, *p*-hydroxyphenylacetic acid, *p*-hydroxyphenylacetamide and *p*-hydroxybenzaldehyde were analyzed in a targeted approach using UPLC-triple quad-MS/MS in either negative or positive mode ESI (details in Supplemental Experimental Methods), with standard curves in the range 0.02-25 µM.

In situ-PCR

The third leaf blades from 7 and 10 day old plants were used for *in situ* analysis of gene expression localization following the procedure previously described (Jørgensen et al., 2005; Jørgensen et al., 2011) which is modified from that of Koltai and Bird (2000). Primers were added according to the gene to be tested (primer sequences in Table S3) and the primer specificities were validated by DNA sequencing of the respective PCR products. The parameters for the PCR reaction were: 2 min at 90 ^{0}C, then 35 cycles at 92 ^{0}C for 30s, 60 ^{0}C for 30 s and 72 ^{0}C for 60 s, and finally 30-60 min at 12 ^{0}C. In the PCR reaction dUTP labelled with digoxigenin (DIG) (Roche) was incorporated into the PCR products, and the gene expression was visualized with a DIG specific antibody conjugated with fluorescein isothiocyanate (FITC) (Roche). The sections were finally embedded in anti-fading agent (Citiflour) and examined by a Leica DMR HC fluorescence microscope and photographed with a Leica DC 300F camera. The images presented from experiment 1 were recorded using the FITC/rhodamine (FI/RH) excitation/emission filters, respectively BP 490/15 and BP605/30, to

differentiate chlorophyll autofluorescence from gene expression. In experiment 2 the enhanced GFP filter with excitation filter BP 470/40 and emission filter BP 525/50 was used.

Accession numbers

GSTL sequence data from this article can be found in the GenBank data libraries under accession numbers (Phytozome identifiers in brackets): *Sb*GSTL1: MF770503 (Sobic.002G421200.2), *Sb*GSTL2: MF770504 (Sobic.009G033200.2), *Sb*GSTL3: MF770505 (Sobic.001G412700.1) *Sb*GSTL4: MF770506 (Sobic.001G412800.2). The sequences of the expressed *Sb*DHAR coding sequences were not experimentally verified. The sequences used correspond to those provided by Lallement et al. (2014a) with Phytozome identifier numbers from version 10 of this database: Sb10g008310, Sb09g001700, Sb09g001690.

Author contributions

N.B. and B.L.M. designed the experiments and wrote the paper, N.B. performed the majority of experiments and analyses; in addition E.H.N., C.C., and C.E.O. performed some of the LC-MS analyses and data handling, M.S.M. and C.E.O. carried out and interpreted NMR experiments, and K.J. carried out and analyzed *in situ*-PCR. D.P.D., R.E., C.C., M.S.M. and C.E.O. provided analytical tools and/or specific scientific advice, and all authors have read and commented on the manuscript. Financial support from the Independent Research Fund Denmark and the VILLUM Foundation is gratefully acknowledged.

Acknowledgements

We thank Dr. Lasse Nielsen for providing data for Figure S4 and Dr. Cecilia Blomstedt and Professor Ros Gleadow for providing the initial batch of *tcd1*seed. Lene Dalsten and Dr. Federico Cozzi are gratefully acknowledged for their technical support in respectively cloning/*in situ*-PCR and LC-MS method development.

Supporting Information

Figure S1. Production of *p*-glucosyloxyphenylacetic acid in excised sorghum leaves upon administration of dhurrin.

Figure S2. Metabolism of L-[U-^{14}C]-tyrosine in sorghum plants.

Figure S3. Structure elucidation of GS-pOHPACN.

Figure S4. GS-pOHPACN in the developing sorghum grain.

Figure S5. Subset of sequence alignment and percent identities of GSTL sequences from different plants.

Figure S6. Expression of *GSTL*s visualized by *in situ* PCR.

Figure S7. Original images from Figure 6.

Table S1. Complete ^{1}H NMR and ^{13}C chemical shift assignments for purified GS-pOHPACN.

Table S2. Turnover numbers of *Arabidopsis* and wheat GSTLs' activities towards GS-pOHPACN.

Table S3. Primers for cloning *GSTL*s from sorghum cDNA and for *in situ-PCR* analysis of *GSTL* and *NIT4* expression.

Methods S1. UHPLC-triple quadrupole-MS/MS analyses of assays with plant protein

Data S1. Full sequence alignment of GSTLs from different plants.

Data S2. Full sequence alignment of GSTLs from different plants, including *At*DHAR1 as outgroup.

References

Abe, H., Uchiyama, M., and Sato, R. (1974). Isolation of phenylacetic acid and its para-hydroxy derivative as auxin-like substances from *Undaria Pinnatifida*. Agr Biol Chem **38**:897-898.

Adewusi, S.R.A. (1990). Turnover of dhurrin in green sorghum seedlings. Plant Physiol **94**:1219-1224.

Agerbirk, N., Warwick, S.I., Hansen, P.R., and Olsen, C.E. (2008). *Sinapis* phylogeny and evolution of glucosinolates and specific nitrile degrading enzymes. Phytochemistry **69**:2937-2949.

Blomstedt, C.K., Gleadow, R.M., O'Donnell, N., Naur, P., Jensen, K., and Laursen, T. (2012). A combined biochemical screen and TILLING approach identifies mutations in *Sorghum bicolor* L. Moench resulting in acyanogenic forage production. Plant Biotechnol J **10**:54-66.

Blomstedt, C.K., O'Donnell, N.H., Bjarnholt, N., Neale, A.D., Hamill, J.D., Møller, B.L., and Gleadow, R.M. (2016). Metabolic consequences of knocking out UGT85B1, the gene encoding the glucosyltransferase required for synthesis of dhurrin in *Sorghum bicolor* (L. Moench). Plant Cell Physiol. **57:**373-386.

Board, P.G., Coggan, M., Cappello, J., Zhou, H., Oakley, A.J., and Anders, M.W. (2008). *S*-(4-Nitrophenacyl)glutathione is a specific substrate for glutathione transferase omega 1-1. Anal Biochem **374**:25-30.

Brazier-Hicks, M., Offen, W.A., Gershater, M.C., Revett, T.J., Lim, E.-K., Bowles, D.J., Davies, G.J., and Edwards, R. (2007). Characterization and engineering of the bifunctional *N*- and *O*-glucosyltransferase involved in xenobiotic metabolism in plants. PNAS **104**:20238-20243.

Busk, P.K., and Møller, B.L. (2002). Dhurrin synthesis in sorghum is regulated at the transcriptional level and induced by nitrogen fertilization in older plants. Plant Physiol. **129**:1222-1231.

Chen, K., Dugas, T.R., and Cole, R.B. (2008). Liquid chromatography-electrospray tandem mass spectrometry investigations of fragmentation pathways of biliary 4,4′-methylenedianiline conjugates produced in rats. Anal Bioanal Chem **391**:271.

Chi, Y.H., Cheng, Y., Vanitha, J., Kumar, N., Ramamoorthy, R., Ramachandran, S., and Jiang, S.Y. (2011). Expansion mechanisms and functional divergence of the glutathione S-transferase family in sorghum and other higher plants. DNA Res **18**:1-16.

Crawford-Sidebotham, T.J. (1972). Role of slugs and snails in maintenance of cyanogenesis polymorphisms of *Lotus corniculatus* and *Trifolium repens*. Heredity **28**:405-411.

Crush, J.R., and Caradus, J.R. (1995). Cyanogenesis potential and iodine concentration in white clover (*Trifolium repens* L) cultivars. New Zeal J Agr Res **38**:309-316.

Cummins, I., Dixon, D.P., Freitag-Pohl, S., Skipsey, M., and Edwards, R. (2011). Multiple roles for plant glutathione transferases in xenobiotic detoxification. Drug Met Rev **43**:266-280.

Ding, L., Qin, S., Li, F., Chi, X., and Laatsch, H. (2008). Isolation, antimicrobial activity, and metabolites of fungus *Cladosporium* sp. associated with red alga *Porphyra yezoensis*. Curr Microbiol **56**:229-235.

Dixon, D.P., Davis, B.G., and Edwards, R. (2002). Functional divergence in the glutathione transferase superfamily in plants - Identification of two classes with putative functions in redox homeostasis in *Arabidopsis thaliana*. J Biol Chem **277**:30859-30869.

Dixon, D.P., and Edwards, R. (2010a). Glutathione transferases. The Arabidopsis Book **8**:e0131. doi.org/0110.1199/tab.0131.

Dixon, D.P., and Edwards, R. (2010b). Roles for stress-inducible Lambda glutathione transferases in flavonoid metabolism in plants as identified by ligand fishing. J Biol Chem **285**:36322-36329.

Dixon, D.P., Hawkins, T., Hussey, P.J., and Edwards, R. (2009). Enzyme activities and subcellular localization of members of the *Arabidopsis* glutathione transferase superfamily. J Exp Bot **60**:1207-1218.

Dixon, D.P., Skipsey, M., and Edwards, R. (2010). Roles for glutathione transferases in plant secondary metabolism. Phytochemistry **71**:338-350.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32**:1792-1797.

Edwards, R., and Dixon, D.P. (2005). Plant glutathione transferases. SAN DIEGO: ELSEVIER ACADEMIC PRESS INC.

Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000). Predicting subcellular localization of proteins based on their *N*-terminal amino acid sequence. J Mol Biol **300**:1005-1016.

Fedrizzi, B., Guella, G., Perenzoni, D., Gasperotti, M., Masuero, D., Vrhovsek, U., and Mattivi, F. (2012). Identification of intermediates involved in the biosynthetic pathway of 3-mercaptohexan-1-ol conjugates in yellow passion fruit (*Passiflora edulis f. flavicarpa*). Phytochemistry **77**:287-293.

Forslund, K., Morant, M., Jørgensen, B., Olsen, C.E., Asamizu, E., Sato, S., Tabata, S., and Bak, S. (2004). Biosynthesis of the nitrile glucosides rhodiocyanoside A and D and the cyanogenic glucosides lotaustralin and linamarin in *Lotus japonicus*. Plant Physiol **135**:71-84.

Fries, L., and Iwasaki, H. (1976). *p*-Hydroxyphenylacetic acid and other phenolic compounds as growth stimulators of red alga *Porphyra tenera*. Plant Sci Lett **6**:299-307.

Geu-Flores, F., Moldrup, M.E., Bottcher, C., Olsen, C.E., Scheel, D., and Halkier, B.A. (2011). Cytosolic gamma-glutamyl peptidases process glutathione conjugates in the biosynthesis of glucosinolates and camalexin in *Arabidopsis*. Plant Cell **23**:2456-2469.

Geu-Flores, F., Nielsen, M.T., Nafisi, M., Moldrup, M.E., Olsen, C.E., Motawia, M.S., and Halkier, B.A. (2009). Glucosinolate engineering identifies gamma-glutamyl peptidase. Nature Chem Biol **5**:575-577.

Gleadow, R.M., and Møller, B.L. (2014). Cyanogenic glycosides: synthesis, physiology, and phenotypic plasticity. Annu Rev Plant Biol **65**:j155-185.

Gleadow, R.M., Veechies, A.C., and Woodrow, I.E. (2003). Cyanogenic *Eucalyptus nobilis* is polymorphic for both prunasin and specific beta-glucosidases. Phytochemistry **63**:699-704.

Gleadow, R.M., and Woodrow, I.E. (2000). Temporal and spatial variation in cyanogenic glycosides in *Eucalyptus cladocalyx*. Tree Physiol **20**:591-598.

Hartmann, T.N., Fricker, M.D., Rennenberg, H., and Meyer, A.J. (2003). Cell-specific measurement of cytosolic glutathione in poplar leaves. Plant Cell Env **26**:965-975.

Hershey, H.P., and Stoner, T.D. (1991). Isolation and characterization of cDNA clones for RNA species induced by substituted benzenesulfonamides in corn. Plant Mol Biol **17**:679-690.

Huang, Y., Xun, R.D., Chen, G.J., and Xun, L.Y. (2008). Maintenance role of a glutathionyl-hydroquinone lyase (PcpF) in pentachlorophenol degradation by *Sphingobium chlorophenolicum* ATCC 39723. J Bacteriol **190**:7595-7600.

Jain, M., Ghanashyam, C., and Bhattacharjee, A. (2010). Comprehensive expression analysis suggests overlapping and specific roles of rice glutathione S-transferase genes during development and stress responses. BMC Genom, doi: doi: 10.1186/1471-2164-11-73.

Jenrich, R., Trompetter, I., Bak, S., Olsen, C.E., Moller, B.L., and Piotrowski, M. (2007). Evolution of heteromeric nitrilase complexes in Poaceae with new functions in nitrile metabolism. PNAS **104**:18848-18853.

Jørgensen, K., Bak, S., Busk, P.K., Sørensen, C., Olsen, C.E., and Puonti-Kaerlas, J. (2005). Cassava plants with a depleted cyanogenic glucoside content in leaves and tubers. Distribution of cyanogenic glucosides, their site of synthesis and transport, and blockage of the biosynthesis by RNA interference technology. Plant Physiol **139**:363-374.

Jørgensen, K., Morant, A.V., Morant, M., Jensen, N.B., Olsen, C.E., Kannangara, R., Motawia, M.S., and Møller, B.L. (2011). Biosynthesis of the cyanogenic glucosides linamarin and lotaustralin in cassava: isolation, biochemical characterization, and expression pattern of CYP71E7, the oxime-metabolizing cytochrome P450 enzyme. Plant Physiol **155**:282-292.

Kobayashi, H., Takase, H., Suzuki, Y., Tanzawa, F., Takata, R., Fujita, K., Kohno, M., Mochizuki, M., Suzuki, S., and Konno, T. (2011). Environmental stress enhances biosynthesis of flavor precursors, S-3-(hexan-1-ol)-glutathione and S-3-(hexan-1-ol)-L-cysteine, in grapevine through glutathione S-transferase activation. J Exp Bot **62**:1325-1336.

Koltai, H., and Bird, D.M. (2000). High throughput cellular localization of specific plant mRNAs by liquid-phase *in situ* reverse transcription-polymerase chain reaction of tissue sections. Plant Physiol **123**:1203-1212.

Korasick, D.A., Enders, T.A., and Strader, L.C. (2013). Auxin biosynthesis and storage forms. J Exp Bot **64**:2541-2555.

Kumar, S., Asif, M.H., Chakrabarty, D., Tripathi, R.D., Dubey, R.S., and Trivedi, P.K. (2013). Differential expression of rice Lambda class GST gene family members during plant growth, development, and in response to stress conditions. Plant Mol Biol Rep **31**:569-580.

Labrou, N.E., Papageorgiou, A.C., Pavli, O., and Flemetakis, E. (2015). Plant GSTome: structure and functional role in xenome network and plant stress response. Curr Opin Biotechnol **32**:186-194.

Lallement, P.-A., Brouwer, B., Keech, O., Hecker, A., and Rouhier, N. (2014a). The still mysterious roles of cysteine-containing glutathione transferases in plants. Front Pharmacol, doi: 10.3389/fphar.2014.00192.

Lallement, P.A., Meux, E., Gualberto, J.M., Prosper, P., Didierjean, C., Saul, F., Haouz, A., Rouhier, N., and Hecker, A. (2014b). Structural and enzymatic insights into Lambda glutathione transferases from *Populus trichocarpa*, monomeric enzymes constituting an early divergent class specific to terrestrial plants. Biochem J **462**:39-52.

Lan, T., Yang, Z.-L., Yang, X., Liu, Y.-J., Wang, X.-R., and Zeng, Q.-Y. (2009). Extensive Functional Diversification of the Populus Glutathione S-Transferase Supergene Family. Plant Cell **21**:3749-3766.

Lieberei, R., Selmar, D., and Biehl, B. (1985). Metabolization of cyanogenic glucosides in *Hevea brasiliensis*. Plant Syst Evol **150:**49-63.

Liu, Y.J., Han, X.M., Ren, L.L., Yang, H.L., and Zeng, Q.Y. (2013). Functional divergence of the glutathione S-transferase supergene family in *Physcomitrella patens* reveals complex patterns of large gene family evolution in land plants. Plant Physiol **161**:773-786.

Mao, C.H., and Anderson, L. (1965). Cyanogenesis in *Sorghum vulgare*. 2. Mechanism of alkaline hydrolysis of dhurrin (*p*-hydroxymandelonitrile glucoside). J Org Chem **30**:603-607.

McCarthy, D.L., Navarrete, S., Willett, W.S., Babbitt, P.C., and Copley, S.D. (1996). Exploration of the relationship between tetrachlorohydroquinone dehalogenase and the glutathione S-transferase superfamily. Biochemistry **35**:14634-14642.

Meux, E., Prosper, P., Ngadin, A., Didierjean, C., Morel, M., Dumarcay, S., Lamant, T., Jacquot, J.P., Favier, F., and Gelhaye, E. (2011). Glutathione transferases of *Phanerochaete chrysosporium* S-glutathionyl-*p*-hydroquinone reductase belongs to a new structural class. J Biol Chem **286**:9162-9173.

Meyer, A.J., May, M.J., and Fricker, M. (2001). Quantitative *in vivo* measurement of glutathione in *Arabidopsis* cells. Plant J **27**:67-78.

Miller, J.M., and Conn, E.E. (1980). Metabolism of hydrogen cyanide by higher plants. Plant Physiol. **65**:1199-1202.

Møller, B.L. (2010). Functional diversifications of cyanogenic glucosides. Curr Opin Plant Biol **13**:338-347.

Møller, B.L., Olsen, C.E., and Motawia, M.S. (2016). General and stereocontrolled approach to the chemical synthesis of naturally occurring cyanogenic glucosides. J Nat Prod **79**:1198-1202.

Neilson, E.H., Goodger, J.Q.D., Motawia, M.S., Bjarnholt, N., Frisch, T., Olsen, C.E., Moller, B.L., and Woodrow, I.E. (2011). Phenylalanine derived cyanogenic diglucosides from *Eucalyptus camphora* and their abundances in relation to ontogeny and tissue type. Phytochemistry **72**:2325-2334.

Neilson, E.H., Goodger, J.Q.D., Woodrow, I.E., and Møller, B.L. (2013). Plant chemical defense: at what cost? Trends Plant Sci **18**:250-258.

Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Prot eng. **10:**1-6.

Nielsen, L.J., Stuart, P., Pičmanová, M., Rasmussen, S., Olsen, C.E., Harholt, J., Møller, B.L., and Bjarnholt, N. (2016). Dhurrin metabolism in the developing grain of Sorghum bicolor (L.) Moench investigated by metabolite profiling and novel clustering analyses of time-resolved transcriptomic data. BMC Genom, doi: 10.1186/s12864-016-3360-4.

Peiser, G.D., Wang, T.T., Hoffman, N.E., Yang, S.F., Liu, H.W., and Walsh, C.T. (1984). Formation of cyanide from carbon-1 of 1-aminocyclopropane-1-carboxylic acid during its conversion to ethylene. PNAS **81**:3059-3063.

Picmanova, M., Neilson, E.H., Motawia, M.S., Olsen, C.E., Agerbirk, N., Gray, C.J., Flitsch, S., Meier, S., Silvestro, D., Jorgensen, K., et al. (2015). A recycling pathway for cyanogenic glycosides evidenced by the comparative metabolic profiling in three cyanogenic plant species. Biochem J **469**:375-389.

Rahantaniaina, M.S., Li, S.C., Chatel-Innocenti, G., Tuzet, A., Issakidis-Bourguet, E., Mhamdi, A., and Noctor, G. (2017). Cytosolic and chloroplastic DHARs cooperate in oxidative stress-driven activation of the salicylic acid pathway. Plant Physiol **174**:956-971**.**

Rezaei, M.K., Shobbar, Z.S., Shahbazi, M., Abedini, R., and Zare, S. (2013). Glutathione S-transferase (GST) family in barley: Identification of members, enzyme activity, and gene expression pattern. J Plant Physiol **170**:1277-1284.

Sanchez-Perez, R., Jørgensen, K., Olsen, C.E., Dicenta, F., and Møller, B.L. (2008). Bitterness in almonds. Plant Physiol **146**:1040-1052.

Seigler, D.S., Pauli, G.F., Frohlich, R., Wegelius, E., Nahrstedt, A., Glander, K.E., and Ebinger, J.E. (2005). Cyanogenic glycosides and menisdaurin from *Guazuma ulmifolia*, *Ostrya virginiana*, *Tiquilia plicata*, and *Tiquilia canescens*. Phytochemistry **66**:1567-1580.

Selmar, D., Irandoost, Z., and Wray, V. (1996). Dhurrin-6'-glucoside, a cyanogenic diglucoside from *Sorghum bicolor*. Phytochemistry **43**:569-572.

Siegien, I., and Bogatek, R. (2006). Cyanide action in plants - from toxic to regulatory. Acta Physiol Plant **28**:483-497.

Simon, S., and Petrášek, J. (2011). Why plants need more than one type of auxin. Plant Sci **180**:454-460.

Su, T.B., Xu, J.A., Li, Y.A., Lei, L., Zhao, L., Yang, H.L., Feng, J.D., Liu, G.Q., and Ren, D.T. (2011). Glutathione-indole-3-acetonitrile is required for camalexin biosynthesis in *Arabidopsis thaliana*. Plant Cell **23**:364-380.

Swain, E., and Poulton, J.E. (1994). Utilization of amygdalin during seedling development of *Prunus serotina*. Plant Physiol **106**:437-445.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol Evol **30**:2725-2729.

Tocheva, E.I., Fortin, P.D., Eltis, L.D., and Murphy, M.E.P. (2006). Structures of ternary complexes of BphK, a bacterial glutathione S-transferase that reductively dechlorinates polychlorinated biphenyl metabolites. J Biol Chem **281**:30933-30940.

Urano, J., Nakagawa, T., Maki, Y., Masumura, T., Tanaka, K., Murata, N., and Ushimaru, T. (2000). Molecular cloning and characterization of a rice dehydroascorbate reductase. Febs Lett **466**:107-111.

Winterberg, B., Du Fall, L.A., Song, X.M., Pascovici, D., Care, N., Molloy, M., Ohms, S., and Solomon, P.S. (2014). The necrotrophic effector protein SnTox3 re-programs metabolism and elicits a strong defence response in susceptible wheat leaves. BMC Plant Biol, doi: 10.1186/s12870-014-0215-5.

Xie, C., Zhong, D., and Chen, X. (2013). A fragmentation-based method for the differentiation of glutathione conjugates by high-resolution mass spectrometry with electrospray ionization. Anal Chim Acta **788**:89-98.

Xun, L.Y., Belchik, S.M., Xun, R., Huang, Y., Zhou, H.N., Sanchez, E., Kang, C., and Board, P.G. (2010). S-Glutathionyl-(chloro)hydroquinone reductases: a novel class of glutathione transferases. Biochem J **428**:419-427.

Zagrobelny, M., Bak, S., and Møller, B.L. (2008). Cyanogenesis in plants and arthropods. Phytochemistry **69**:1457-1468.

Tables

Table 1. Kinetic parameters determined for GSTL enzymes from sorghum compared to other plants.

	Kinetic parameters					
	k_{cat} (min^{-1} ± SEM)†		K_m (μM ± SEM)†		k_{cat}/K_m (min^{-1} μM^{-1} ± SEM)	
	GS-pOHPACN‡		GS-pOHPACN		GS-pOHPACN	
Enzyme	1	2	1	2	1	2
*Sb*GSTL1	203 ± 12	674 ± 37	30 ± 5	25 ± 4	7 ± 1	28 ± 5
*Sb*GSTL2	143 ± 8	290 ± 21	4 ± 1	4 ± 1	40 ± 9	73 ± 25
*Sb*GSTL3			n.a.§			
*Sb*GSTL4	n.d.¶	14 ± 1	n.d.	10 ± 2		1 ± 0.3
*Sb*ΔGSTL4	8 ± 0.4	7 ± 0.3	15 ± 2	10 ± 2	0.6 ± 0.1	0.7 ± 0.1
	Turnover number# (min^{-1}; SD)					
	GS-pOHPACN					
Enzyme	1	2				
*Ta*GSTL1	n.a.	n.a.				
*Ta*GSTL3	56; 7	64; 6				
*At*GSTL3	15; 0.3	14; 0.5				

†Results for k_{cat} and K_m are means of three experimental replicates.
‡GS-pOHPACN 1 and 2 refer to the earlier and later eluting epimer, respectively.
§n.a. = no activity
¶n.d. = not determined (due to low yield of purified *Sb*GSTL4, k_{cat} and K_m were only determined for GS-pOHPACN 2)
#Due to limited substrate availability, *Ta* and *At* enzymes were assayed at one fixed concentration, and only the two most active and the completely inactive *Ta*GSTL1 were assayed in duplicate, with the means of the resulting turnover numbers shown here. Single measurements of the activity of the remaining *Ta* and *At* enzymes are found in Table S2.

Figure legends

Figure 1. Bioactivation, detoxification and putative recycling pathway of the cyanogenic glucoside dhurrin in sorghum. A. Dhurrin is hydrolyzed by a specific β-glucosidase, dhurrinase (DHR), forming an unstable cyanohydrin, p-hydroxymandelonitrile, which releases hydrogen cyanide (HCN), either spontaneously or mediated by the enzyme hydroxynitrile lyase (HNL). B. β-Cyanoalanine synthase (CAS) incorporates HCN into β-cyanoalanine that can be converted to asparagine, aspartate and ammonia by heteromers of nitrilases (NIT4) of the A and B types. C. Glutathione (GSH) replaces the glucose moiety in dhurrin, either spontaneously or mediated by an unknown glutathione transferase (GST). The resulting GS-conjugate of *p*-hydroxyphenyl acetonitrile (GS-pOHPACN, *p*-hydroxyphenyl(S-glutathione)acetonitrile) is cleaved by GSTs of the lambda class (GSTL) converting reduced GSH to its oxidized form (GSSG) and releasing *p*-hydroxyphenyl acetonitrile. This is the substrate for the sorghum NIT4 heteromer which hydrolyses *p*-hydroxyphenyl acetonitrile to *p*-hydroxyphenylacetic acid and free ammonia, available for primary/general metabolism.

Figure 2. Content of dhurrin and p-glucosyloxyphenylacetic acid in sorghum plants of increasing age. pGPAAc = *p*-glucosyloxyphenylacetic acid. One biological replicate is the aerial tissue of a whole plant, including mesocotyl, node, coleoptile, sheaths and leaves. Each point is an average of three plants, except the last three points where 1-3 additional plants were analyzed to confirm that the large standard deviations were representative of the biological variation in the sample material. A. Total content of each compound plotted as function of plant age. Error bars represent standard deviations of biological replicates. B. The average of ratios between the two compounds in each replicate.

Figure 3. *In planta* production of p-glucosyloxyphenylacetic acid from dhurrin. A. The second- and third-oldest leaves from 7 days old plants were excised below the blade and administered either 500 nmol dhurrin in 50 µl water ('dhurrin') or the same volume of water ('water') by uptake through the cut. After uptake, the leaves were incubated for 5 h and then snap frozen. 'Control' leaves were harvested directly into liquid nitrogen at t = 0. "Parent" = parent line, "tcd1" = "total cyanide deficient" mutant line. Error bars represent standard deviations of biological replicates, n = 3. B. Production of ^{14}C-labelled dhurrin and *p*-glucosyloxyphenylacetic acid in WT plants following administration of ^{14}C-tyrosine to intact plants. The results were analyzed by TLC, here shown for extracts of leaf 2 (TLC analysis of total plant extracts in Figure S2). 1 = *p*-hydroxybenzaldehyde, 2 = dhurrin, 3 = *p*-glucosyloxyphenylacetic acid. C. Quantification of ^{14}C in relevant fractions from the experiment described in B.

Figure 4. Dhurrin metabolizing enzyme activities in sorghum plants. Soluble protein was extracted from plants of increasing age (x-axes, 5-7-10-13-17-24 days after sowing) and incubated with different substrates from the proposed pathway for dhurrin turnover. A. Substrate = 0.2 mM GS-pOHPACN 1, incubation = 5 min; B. Substrate = 0.2 mM GS-pOHPACN 2, incubation = 5 min; C. Substrate = 2.5 mM dhurrin, incubation = 45 min. Due to order of magnitude difference in recovery of different products the results have been split into C1: Substrate recovery (dhurrin) and

bioactivation pathway (*p*-hydroxybenzaldehyde), and C2: Recycling pathway products. Note the different scale and hatched y-axis. Detected products across all charts: Dark grey = GS-pOHPACN, light grey = *p*-hydroxyphenyacetonitrile, white = *p*-hydroxyphenylacetic acid, black = dhurrin, hatched bars = *p*-hydroxybenzaldehyde. In all charts the products are presented as % recovery of added substrate; in the control the substrate was incubated in assay buffer without protein. All experiments were performed in triplicate (technical replicates), and error bars represent standard deviations of the sum of quantified products.

Figure 5. Maximum likelihood phylogenetic tree of GSTLs from different plants. Sequences from other plants than sorghum: *Arabidopsis thaliana, At*GSTL1-*At*GSTL3 (Dixon et al. 2002; Dixon et al. 2009), wheat (*Triticum aestivum*), *Ta*GSTL1-*Ta*GSTL3 (Dixon and Edwards 2010b), rice (*Oryza sativa*) *Os*GSTL1-*Os*GSTL3 (Kumar et al. 2013), maize (*Zea mays*), *Zm* In2-1 (Hershey and Stoner 1991), poplar (*Populus trichocarpa*), *Pt*GSTL1-*Pt*GSTL3 (Lan et al. 2009). Bootstrap = 1000. Condensed branch lengths were computed with 70 as cut off value. Due to significant deviations at the ends of some of the proteins (*Sb*GSTL4, *Os*GSTL3 and *Pt*GSTL1-*Pt*GSTL3 (Lallement et al., 2014b)) from the remaining sequences, another tree was constructed where the ends were trimmed and the sequences realigned. The result was the same.

Figure 6. Expression of GSTLs and NIT4s visualized by *in situ* PCR. Gene expression is visualized as green fluorescence in all images. In some cases, fluorescence pertaining from expression of the genes of interest was challenged by strong chlorophyll and cell wall autofluorescence. Images were therefore captured with two different filters to select the settings that best distinguish between the different sources of fluorescence. In experiment 1 the best images were recorded using the Leica FITC/rhodamine (FI/RH) filter to differentiate chlorophyll fluorescence (magenta, false color). In experiment 2 the best images were captured with the Leica enhanced GFP filter, providing a stronger green fluorescence from gene expression. chl = chlorophyll, xyl = xylem vessels. Experiment 1: A. *SbGSTL1*, day 10. B. *SbGSTL2*, day 10. C. *SbGSTL4*, day 7. D. Negative control, day 10. The green xylem cell wall autofluorescence is visualized in the negative control in 1.D. This autofluorescence is also strong in images 1.A and 1.C. The green structures in the upper right corner of 1.A and lower right corner of 1.C are caused by autofluorescence of the mid rib cell walls. In 1.A-1.C green fluorescence inside cells that are not xylem vessels corresponds to expression of the respective genes. Experiment 2, all from day 10: A. *SbGSTL1*. B. *SbNIT4A*. C. *SbNIT4B1*. D. *SbNIT4B2*. Image 2.C is similar to the negative control of experiment 2 (Figure S6), with only xylem and epidermal cell wall autofluorescence and a weak contribution from chlorophyll visible; in the remaining images green fluorescence not pertaining from epidermal cell walls corresponds mainly to gene expression with only a minor contribution from xylem cell walls. In images from experiment 1, false colors have been applied to avoid red/green comparison. In images from experiment 2, the contrast has been increased to improve visualization. All original images can be viewed in Figure S7.

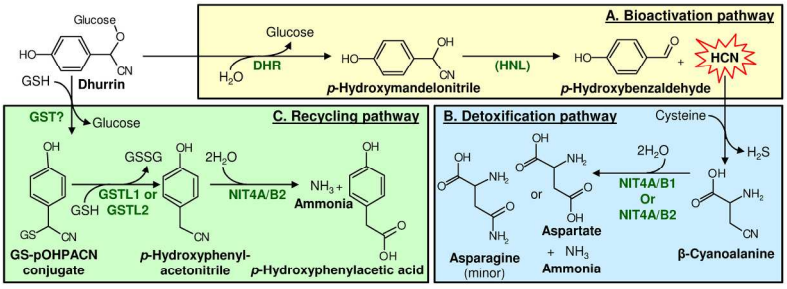
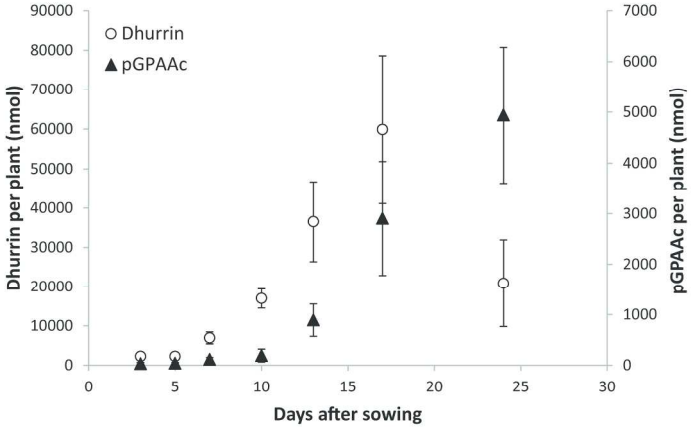


Figure 1. Bioactivation, detoxification and putative recycling pathway of the cyanogenic glucoside dhurrin in sorghum.
A. Dhurrin is hydrolyzed by a specific β-glucosidase, dhurrinase (DHR), forming an unstable cyanohydrin, *p*-hydroxymandelonitrile, which releases hydrogen cyanide (HCN), either spontaneously or mediated by the enzyme hydroxynitrile lyase (HNL). B. β-Cyanoalanine synthase (CAS) incorporates HCN into β-cyanoalanine that can be converted to asparagine, aspartate and ammonia by heteromers of nitrilases (NIT4) of the A and B types. C. Glutathione (GSH) replaces the glucose moiety in dhurrin, either spontaneously or mediated by an unknown glutathione transferase (GST). The resulting GS-conjugate of *p*-hydroxyphenylacetonitrile (GS-pOHPACN, *p*-hydroxyphenyl(S-glutathione)acetonitrile) is cleaved by GSTs of the lambda class (GSTL) converting reduced GSH to its oxidized form (GSSG) and releasing *p*-hydroxyphenyl acetonitrile. This is the substrate for the sorghum NIT4 heteromer which hydrolyses *p*-hydroxyphenylacetonitrile to *p*-hydroxyphenylacetic acid and free ammonia, available for primary/general metabolism.

209x148mm (300 x 300 DPI)

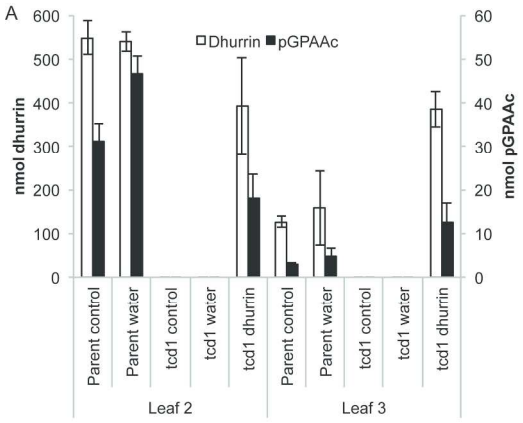


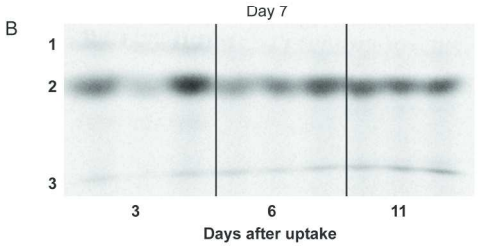
Days†	Dhurrin:pGPAAc Mean; SD
3	149; 62
5	117; 78
7	155; 68
10	351; 353
13	82; 30
17	44; 11
24	8; 4

†Days after sowing

Figure 2. Content and ratio of dhurrin and p-glucosyloxyphenylacetic acid in sorghum plants of increasing age. pGPAAc = *p*-glucosyloxyphenylacetic acid. One biological replicate is the aerial tissue of a whole plant, including mesocotyl, node, coleoptile, sheaths and leaves. Each point is an average of three plants, except the last three points where 1-3 additional plants were analyzed to confirm that the large standard deviations were representative of the biological variation in the sample material. A. Total content of each compound plotted as function of plant age. Error bars represent standard deviations of biological replicates. B. The average of ratios between the two compounds in each replicate.

209x123mm (300 x 300 DPI)





C

Relative amount (s.d.)			
Days after uptake	3	6	11
Total ^{14}C†	100 (20)	82.5 (2.5)	62 (12)
Total ^{14}C-dhurrin‡	100 (20)	66 (8)	39 (9)
^{14}C-pGPAAc (leaf)‡	100 (45)	126 (19)	233 (78)
^{14}C-dhurrin:^{14}C-pGPAAc (leaf)‡	15 (5.5)	9 (3)	5 (1.2)

†Determined in total plant extracts by scintillation counting
‡Determined by spot quantification in TLC analyses of respectively total plant extracts (Figure S2) and leaf 2 extracts (leaf) (panel B).

Figure 3. In planta production of *p*-glucosyloxyphenylacetic acid from dhurrin.
A. Administration of dhurrin to leaves from EMS mutant plants. The second- and third-oldest leaves from 7 days old plants were excised below the blade and administered either 500 nmol dhurrin in 50 µl water ('dhurrin') or the same volume of water ('water') by uptake through the cut. After uptake, the leaves were incubated for 5 h and then snap frozen. 'Control' leaves were harvested directly into liquid nitrogen at t = 0. "Parent" = parent line, "tcd1" = "total cyanide deficient" mutant line. Error bars represent standard deviations of biological replicates, n = 3. B. Production of ^{14}C-labelled dhurrin and *p*-glucosyloxyphenylacetic acid in WT plants following administration of ^{14}C-tyrosine to intact plants. The results were analyzed by TLC, here shown for extracts of leaf 2 (TLC analysis of total plant extracts in Figure S2). 1 = p-hydroxybenzaldehyde, 2 = dhurrin, 3 = p-glucosyloxyphenylacetic acid. C. Quantification of ^{14}C in relevant fractions from the experiment described in B.

156x264mm (300 x 300 DPI)

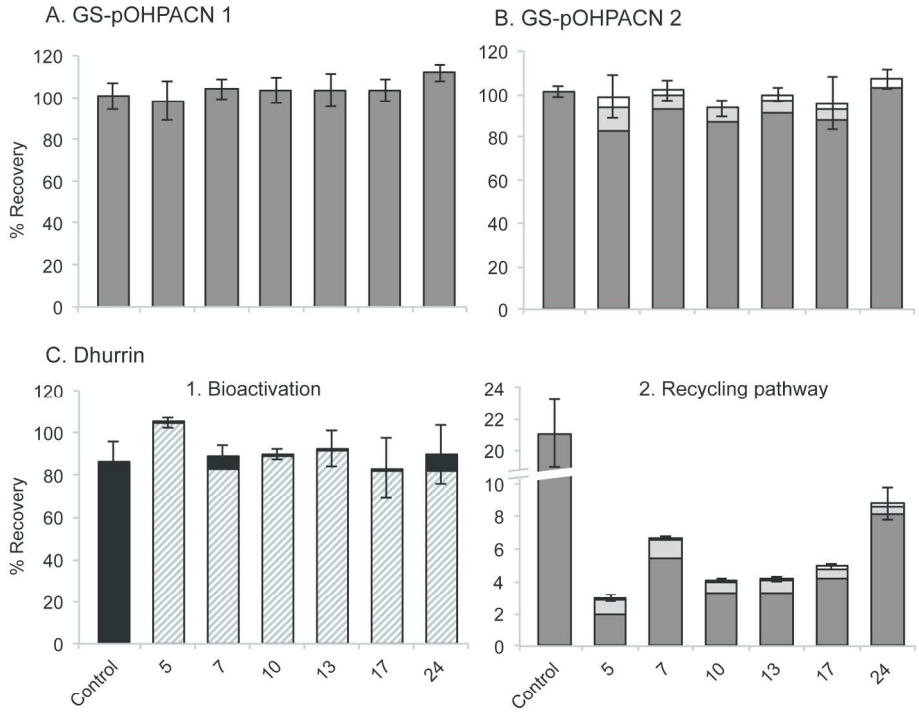


Figure 4. Dhurrin metabolizing enzyme activities in sorghum plants.
Soluble protein was extracted from plants of increasing age (x-axes, 5-7-10-13-17-24 days after sowing) and incubated with different substrates from the proposed pathway for dhurrin recycling. A. Substrate = 0.2 mM GS-pOHPACN 1, incubation = 5 min; B. Substrate = 0.2 mM GS-pOHPACN 2, incubation = 5 min; C. Substrate = 2.5 mM dhurrin, incubation = 45 min. Due to order of magnitude difference in recovery of different products the results have been split into C1: Substrate recovery (dhurrin) and bioactivation pathway (p-hydroxybenzaldehyde), and C2: Recycling pathway products. Note the different scale and hatched y-axis. Detected products across all charts: Dark grey = GS-pOHPACN, light grey = p-hydroxyphenyacetonitrile, white = p-hydroxyphenylacetic acid, black = dhurrin, hatched bars = p-hydroxybenzaldehyde. In all diagrams the products are presented as % recovery of added substrate; in the control, the substrate was incubated in assay buffer including GSH, without protein. All experiments were performed in triplicate (technical replicates), and error bars represent standard deviations of the sum of quantified products.

194x191mm (300 x 300 DPI)

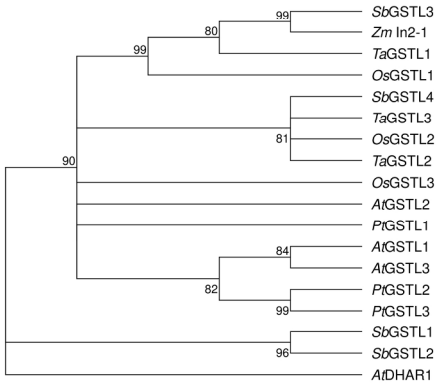


Figure 5. Maximum likelihood phylogenetic tree of GSTLs from different plants. Sequences from other plants than sorghum: *Arabidopsis thaliana*, *At*GSTL1-*At*GSTL3 (Dixon et al. 2002; Dixon et al. 2009), wheat (*Triticum aestivum*), *Ta*GSTL1-*Ta*GSTL3 (Dixon and Edwards 2010b), rice (*Oryza sativa*) *Os*GSTL1-*Os*GSTL3 (Kumar et al. 2013), maize (*Zea mays*), *Zm* In2-1 (Hershey and Stoner 1991), poplar (*Populus trichocarpa*), *Pt*GSTL1-*Pt*GSTL3 (Lan et al. 2009). Bootstrap = 1000. Condensed branch lengths were computed with 75 as cut off value. Due to significant deviations at the ends of some of the proteins (*Sb*GSTL4, *Os*GSTL3 and *Pt*GSTL1-*Pt*GSTL3 (Lallement et al., 2014b)) from the remaining sequences, another tree was constructed where the ends were trimmed and the sequences realigned. The result was the same.

209x148mm (300 x 300 DPI)

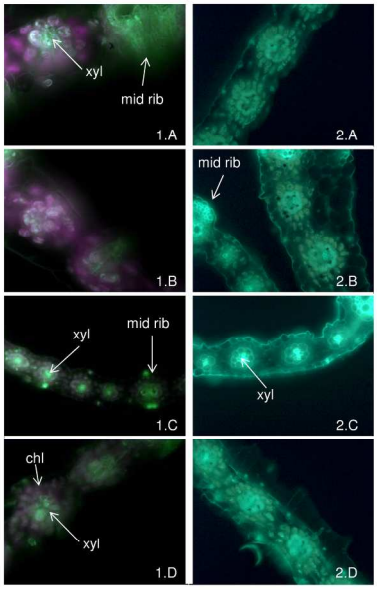


Figure 6. Expression of GSTLs and NIT4s visualized by *in situ* PCR. Gene expression is visualized as green fluorescence in all images. In some cases, fluorescence pertaining from expression of the genes of interest was challenged by strong chlorophyll and cell wall autofluorescence. Images were therefore captured with two different filters to select the settings that best distinguish between the different sources of fluorescence. In experiment 1 the best images were recorded using the Leica FITC/rhodamine (FI/RH) filter to differentiate chlorophyll fluorescence (magenta, false color). In experiment 2 the best images were captured with the Leica enhanced GFP filter, providing a stronger green fluorescence from gene expression. chl = chlorophyll, xyl = xylem vessels. Experiment 1: A. *SbGSTL1*, day 10. B. *SbGSTL2*, day 10. C. *SbGSTL4*, day 7. D. Negative control, day 10. The green xylem cell wall autofluorescence is visualized in the negative control in 1.D. This autofluorescence is also strong in images 1.A and 1.C. The green structures in the upper right corner of 1.A and lower right corner of 1.C are caused by autofluorescence of the mid rib cell walls. In 1.A-1.C green fluorescence inside cells that are not xylem vessels corresponds to expression of the respective genes. Experiment 2, all from day 10: A. *SbGSTL1*. B. *SbNIT4A*. C. *SbNIT4B1*. D. *SbNIT4B2*. Image 2.C is similar to the negative control of experiment 2 (Figure S6), with only xylem and epidermal cell wall autofluorescence and a weak contribution from chlorophyll visible; in the remaining images green fluorescence not pertaining from epidermal cell walls corresponds mainly to gene expression with only a minor contribution from xylem cell walls. In images from experiment 1, false colors have been applied to avoid red/green comparison. In images from experiment 2, the contrast has been increased to improve visualization. All original images can be viewed in Figure S7.

209x148mm (300 x 300 DPI)

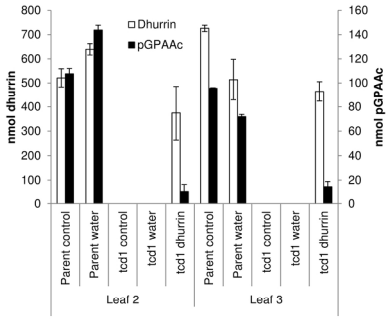


Figure S1. Production of *p*-glucosyloxyphenylacetic acid in excised sorghum leaves upon administration of dhurrin. This is the same experiment as in Figure 3 (refer to this for detailed legend), but leaves were taken from 14 days old plants. "Parent" = parent line, "*tcd1*" = "total cyanide deficient" mutant line, "control" = harvested directly into liquid nitrogen, "water" = leaves administered water, "dhurrin" = leaves administered 500 nmol dhurrin.

209x148mm (300 x 300 DPI)

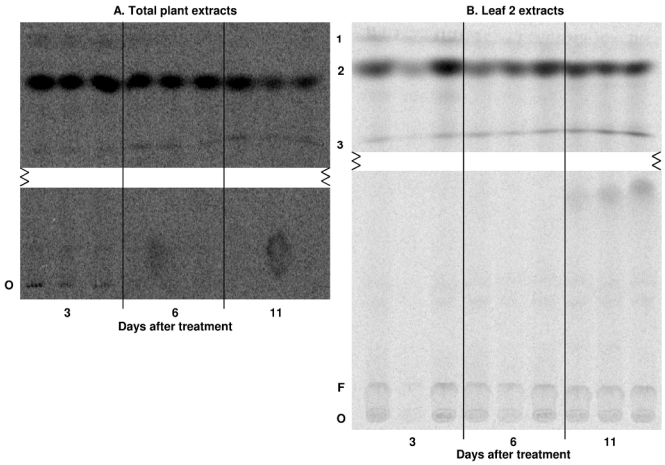


Figure S2. Metabolism of L-[U-^{14}C]-tyrosine in sorghum plants.
Additional data from the experiment shown in Figure 3B+C. In the present figure all areas of TLC plates displaying radiolabelled spots/bands are shown. In both A and B, blank sections of the plates have been omited. 1 = *p*-hydroxybenzaldehyde, 2 = dhurrin, 3 = *p*-glucosyloxyphenylacetic acid, O = origin, F = Focused origin following eution in 100 % methanol. A. TLC analysis of total plant extracts (all tissue extracts combined). In the displayed analysis, the threshold intensity was decreased to allow *p*-glucosyloxyphenylacetic acid spots to be visible, causing over saturation of dhurrin spots. The quantification of relative intensities of dhurrin spots shown in the table in Figure 3C were obtained using a higher threshold intensity to decrease the intensity of dhurrin spots. B. TLC analysis of extracts of leaf 2, displaying all areas of the plate where there was visible radiolabel.

209x148mm (300 x 300 DPI)

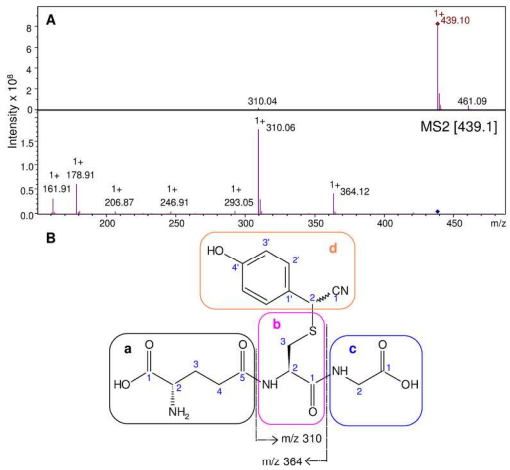


Figure S3. Structure elucidation of GS-pOHPACN.. A. LC-MS/MS analysis of GS-pOHPACN gave rise to peaks of m/z 439 and m/z 461, corresponding respectively to the hypothetical $[GS\text{-}pOHPACN+H]^+$ and $[GS\text{-}pOHPACN+Na]^+$ adduct ions. Parent ion m/z 439 (top pane) gave rise to fragments at m/z 310 and m/z 364 (bottom pane), corresponding to cleavage of the peptide bonds (Chen et al. 2008; Xie et al. 2013) indicated in B. Structure of GS-pOHPACN. The chemical shift assignments of the a, b, c and d substructures of the two epimers are given in Table S1. Structures a-c are individual parts of the GSH skeleton while structure d is the aglycone of *(S)*-dhurrin and/or its epimer *(R)*-taxiphyllin.

209x148mm (300 x 300 DPI)

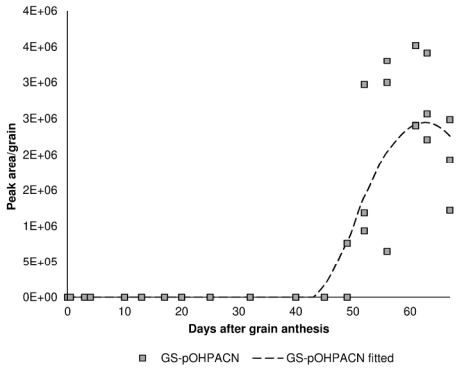
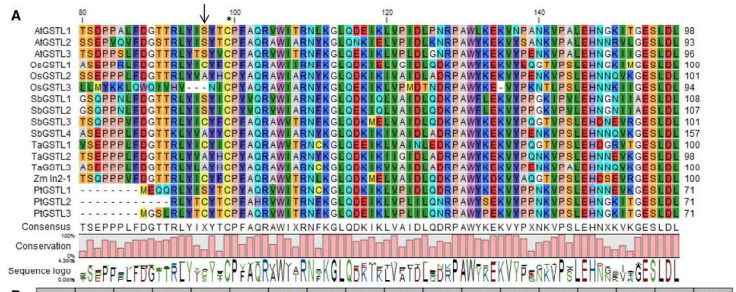


Figure S4. GS-pOHPACN in the developing sorghum grain. The data were extracted from the LC-MS analyses of Nielsen et al. (2016) where the dhurrin content peaked around day 25 in developing sorghum grain. No GS-pOHPACN standard curve was included in this study and therefore the data points are shown as peak area/grain. For comparison, the peak area per grain at maximum dhurrin content was approximately $1x10^8$ (Nielsen et al., 2016), meaning that in the developing grain the amount of GS-pOHPACN is comparatively low, similar to the situation in the young plants (Figure 2 + data in text). Each data point represents a biological replicate of extraction of one grain (three per time point), and the procedures for extraction, analyses and curve fitting are described by Nielsen et al. (2016).

209x148mm (300 x 300 DPI)



B

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
AtGSTL1	1		54,51	72,38	53,47	54,25	36,67	43,61	44,15	49,80	46,33	50,82	55,33	57,38	48,36	56,43	63,33	59,34
AtGSTL2	2	54,51		53,91	53,04	61,29	41,47	42,64	43,56	52,02	46,05	51,42	55,47	54,66	50,61	58,82	52,72	51,26
AtGSTL3	3	72,38	53,91		55,74	55,69	39,13	47,17	48,11	51,64	48,16	51,85	57,26	56,79	49,79	56,30	64,56	60,92
OsGSTL1	4	53,47	53,04	55,74		68,29	40,53	48,68	49,62	78,78	56,48	77,05	66,39	64,75	77,05	51,64	50,61	48,77
OsGSTL2	5	54,25	61,29	55,69	68,29		47,02	50,38	50,94	66,40	67,55	66,26	84,96	77,24	63,41	56,10	52,63	50,00
OsGSTL3	6	36,67	41,47	39,13	40,53	47,02		33,65	34,07	37,09	37,71	38,21	44,85	43,19	36,21	39,80	36,67	34,78
SbGSTL1	7	43,61	42,64	47,17	48,68	50,38	33,65		94,72	48,30	41,27	47,55	49,06	50,19	46,79	41,13	42,26	41,51
SbGSTL2	8	44,15	43,56	48,11	49,62	50,94	34,07	94,72		48,86	41,90	48,48	50,00	50,76	46,59	41,29	41,57	40,91
SbGSTL3	9	49,80	52,02	51,64	78,78	66,40	37,09	48,30	48,86		54,15	78,28	65,57	63,93	89,75	50,61	47,76	44,90
SbGSTL4	10	46,33	46,05	48,16	56,48	67,55	37,71	41,27	41,90	54,15		53,00	67,67	63,33	52,00	45,51	45,00	42,52
TaGSTL1	11	50,82	51,42	51,85	77,05	66,26	38,21	47,55	48,48	78,28	53,00		64,20	63,37	77,79	49,59	49,77	46,72
TaGSTL2	12	55,33	55,47	57,26	66,39	84,96	44,85	49,06	50,00	65,57	67,67	64,20		79,42	63,37	57,02	53,72	50,83
TaGSTL3	13	57,38	54,66	56,79	64,75	77,24	43,19	50,19	50,76	63,93	63,33	63,37	79,42		62,14	54,51	51,23	48,36
Zm In2-1	14	48,36	50,61	49,79	77,05	63,41	36,21	46,79	46,59	89,75	52,00	77,78	63,37	62,14		48,77	45,49	43,85
PtGSTL1	15	56,43	58,82	56,30	51,64	56,10	39,80	41,13	41,29	50,61	45,51	49,59	57,02	54,51	48,77		68,81	68,54
PtGSTL2	16	63,33	52,72	64,56	50,61	52,63	36,67	42,26	41,67	47,76	45,00	48,77	53,72	51,23	45,49	68,81		82,57
PtGSTL3	17	59,34	51,26	60,92	48,77	50,00	34,78	41,51	40,91	44,90	42,52	46,72	50,83	48,36	43,85	68,54	82,57	

Figure S5. Subset of sequence alignment and percent identities of GSTL sequences from different plants. Both A and B are based on the full protein sequence alignment which can be downloaded from Supporting Data S1. A. Subset of alignment showing the conserved catalytic cysteine residue (*) and the additional cysteine residue in 6 sequences (arrow). B. Percent identity matrix. For plant name abbreviations and references refer to the legend of Figure 5.

209x148mm (300 x 300 DPI)

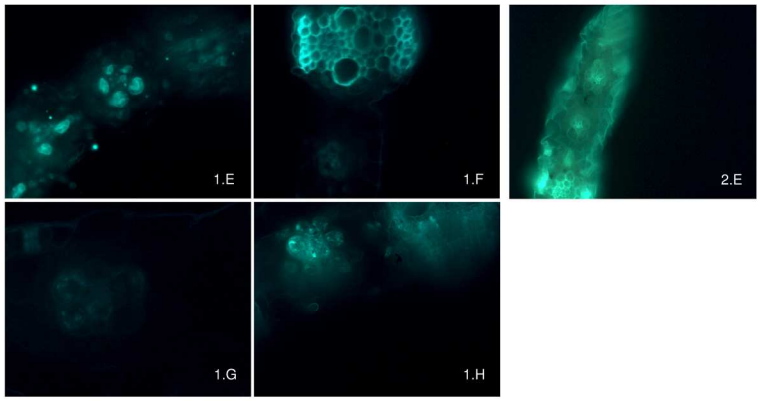


Figure S6. Expression of *GSTLs* visualized by *in situ* PCR. Images 1.E-1.H originate from experiment 1 in Figure 6; 2.E from experiment 2. All images were captured with the Leica enhanced GFP filter. Experiment 1: E. *SbGSTL1*, day 7. F. *SbGSTLA*, day 10. G. Negative control, day 10. H. Same as Figure 7, 1.A (*SbGSTL1*, day 10), but captured with the Leica enhanced GFP filter for comparison with 1.E-1.G. Experiment 2: Negative control, day 10. In image 1.E the fluorescence corresponding to expression of *SbGSTL1* mainly in sheath bundle cells is much stronger than cell wall and chlorophyll autofluorescence, whereas in images 1.F, 1.G and 2.E contributions from these sources cause the only visible fluorescence, with 1.F and 1.G in particular showing the characteristic chlorofyll rich chloroplasts in the bundle sheath cells of C4 plants.

209x148mm (300 x 300 DPI)

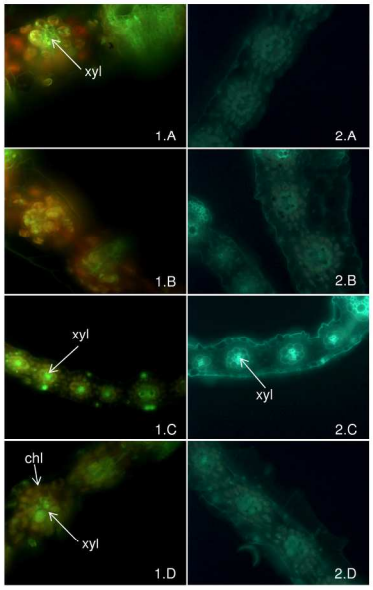


Figure S7. Original images from Figure 6. See Figure 6 for legend. The red color in experiment 1 of the present figure has been changed to magenta in Figure 6.

209x148mm (300 x 300 DPI)

Table S1. Complete ^{1}H NMR and ^{13}C chemical shift assignments for purified GS-pOHPACN.

C	GS-pOHPACN 1 ^{1}H (J)†	^{13}C†	GS-pOHPACN 2 ^{1}H (J)	^{13}C
		Sub-structure a		
1	-	174.4	-	174.6
2	3.78 (t, J = 6.4)	54.8	3.77 (t, J = 6.4)	54.9
3	2.13 (m, $J_{2,3}$ = 7.5, $J_{3,4}$ = 7.2)	26.9	2.14 (q, J =7.5)	27.0
4	2.47 (t, J = 7.8)	32.1	2.50 (m)	32.2
5	-	175.4	-	175.6
		Sub-structure b		
1	-	172.7	-	172.6
2	4.53 (dd, J = 7.3)	53.4	4.61 (dd, $J_{2,3b}$ =5.1, $J_{2,3a}$ = 8.8)	53.3
3	3.14 (dd, $J_{2,3a}$ = 7.5, $J_{3a,3b}$ =14.3, H3-a) 3.18 (dd, $J_{2,3a}$ =5.8, $J_{3a,3b}$ =14.3, H3-b)	33.4	2.96 (dd, $J_{2,3a}$ = 8.8, $J_{3a,3b}$ = 14.3, H3-a) 3.29 (dd, $J_{2,3a}$ =5.1, $J_{3a,3b}$ =14.3, H3-b)	33.7
		Sub-structure c		
1	-	174.9	-	175.3
2	3.90 (s)	42.8	3.85 (d, J_{gem} = 17.5, H2-a) 3.90 (d, J_{gem} = 17.5, H2-b)	43.1
		Sub-structure d		
1	-	120.1	-	120.1
2	5.25 (s)	36.8	5.24 (s)	36.6
1’	-	124.9	-	124.7
2’	7.41 (d, J = 8.6)	130.2	7.40 (d, J = 8.6)	130.2
3’	6.94 (d, J = 8.6)	117.0	6.94 (d, J = 8.6)	117.0
4’	-	157.2		157.2

Sub-structures a-d and numbering of the carbon atoms are outlined in Figure S2. Structures a-c are individual parts of the GSH skeleton while structure d is the aglycone of *(S)*-dhurrin and/or its epimer *(R)*-taxiphyllin.

†The chemical shifts (δ) are given in ppm relative to dioxane as internal reference (δ 3.75 for ^{1}H and δ 67.40 for ^{13}C) ppm, and the coupling constants (J) in Hz. Instrument details can be found in the Methods section.

The multiplicities are reported as follows: s = singlet, d = doublet, dd = doublet of doublets, ddd = doublet of doublet of doublets, m = multiplet, t = triplet.

Table S2. Turnover numbers of single measurements of Arabidopsis and wheat GSTLs' activities towards GS-pOHPACN.

Enzyme	Substrate	Turnover number (min^{-1})
*Ta*GSTL2	GS-pOHPACN 1[†]	9
	GS-pOHPACN 2	7
*At*GSTL1	GS-pOHPACN 1	8
	GS-pOHPACN 2	6
*At*GSTL2	GS-pOHPACN 1	5
	GS-pOHPACN 2	7

[†]GS-pOHPACN 1 and 2 refer to the first and last eluting epimer, respectively.
The results originate from the experiment giving rise to the data in the lower part of Table 1.

Table S3. Primers for cloning *GSTL*s from sorghum cDNA and for *in situ-PCR* analysis of *GSTL* and *NIT4* expression.

Cloning primers		
Gene	Forward	Reverse
GSTL1	ATGGCGACGGCGGCGGCACCCGAT	TCAGCTGAGCTGCTAAGGGGTATGGCGACG
GSTL2	ATGGCGACGGCGGCGGCACCCGAT	CCTCCCCTAGGGATGATAAGATCAGATCAGG CTA
GSTL3	CATGGCCGCAGCCGCAGGA	ACTAATGCCGCTCTGTCACG
GSTL4	CCATGAATTCGCTCGCTTTC	TCATACCCCAAGCCGCTTCT
***In situ*-PCR primers**		
GSTL1	CGCAGTCGACATGGTTTATG	GGGCTTTGTATCCCGAACT
GSTL2	CGCAGTCGACATGGTTTATG	GGGCTTTGTATCCCAAACTT
GSTL4	GCTTGCCTGTAAGCTCCAAG	GAAGGCACCTTGTTTTCTGG
Nit4A	TTACACGATTAGCTGCTATG	ATCTCAATACTTTACCATAC
Nit4B1	ATGTATGCCAAAGGTATTCAG	GACCTGCCAACACTGTTCCG
Nit4B2	TGACCTCGACCTTGGAGAGATT	ATTTGTACACAGAGTCGTCGC

Methods S1. UHPLC-triple quadrupole-MS/MS analyses of assays with plant protein

Chromatography was performed on an Advance UHPLC system (Bruker). Separation was achieved on a Zorbax SB-C18 column (50 x 2.1 mm, 1.8 µm, Agilent Technologies). The mobile phases were A: Water with 0.05 % formic acid, and B: acetonitrile with 0.05 % formic acid. The mobile phase flow rate was 300 $\mu L*min^{-1}$, and the elution profile was as follows: 0-0.5 min, 5 % B; 0.5-2.5 min, 5-100 % B; 2.5-3.5 min 100 % B; 3.5-3.7 min 100-5 % B and 3.7-6.2 min 5 % B. The column temperature was maintained at 40°C. The liquid chromatography was coupled to an EVOQ Elite TripleQuadrupole mass spectrometer (Bruker) equipped with an electrospray ion source (ESI) operated in positive and negative ionization mode. The instrument parameters were optimized by infusion experiments with pure standards. The ion spray voltage was maintained at +4000 V or -4000 V for positive and negative ionization mode respectively. Cone temperature was set to 300 °C and cone gas to 20 psi. Heated probe temperature was set to 300 °C and probe gas flow set to 50 psi. Nebulizing gas was set to 60 psi and collision gas to 1.5 mTorr. Multiple reaction monitoring (MRM) was used to monitor analyte parent ion → product ion transitions. MRMs were determined from direct infusion experiments of reference standards, detected in the Q1 quadrupole as either Na^+ or H^+ adducts in positive mode or as deprotonated ions in negative mode. For each compound one product ion was used as quantifier (^{Q}) and one or two product ions as qualifiers. The MRMs and the corresponding collision energies were as follows. Positive mode: $[Dhurrin+Na]^+$ *m/z* 334.1 → 145.0^Q (-15 eV), 334.1 → 185.0 (-15 eV), 334.1 → 307.0 (-10 eV); $[GS\text{-}pOHPACN+H]^+$ *m/z* 439.1 → 179.0^Q (-17 eV), 439.1 → 310.0 (-11 eV), 439.1 → 364.0 (-12 eV); [*p*-hydroxyphenylacetonitrile-H]$^-$ *m/z* 134.1 → 107.1^Q (-9 eV), 134.1 → 117.1 (-4 eV), 134.1 → 77.2 (-26 eV); [*p*-hydroxyphenylacetamide-H]$^-$ *m/z* 152.1 → 107.1^Q (-14 eV), 152.1 → 77.2 (-32 eV). Negative mode: [*p*-hydroxyphenylacetic acid-H]$^-$, *m/z* 151.1 → 107.1^Q (+6 eV), 151.1 → 92.2 (+6 eV); [*p*-hydroxybenzaldehyde-H]$^-$ *m/z* 121.1 → 92.2^Q (+23 eV) and 121.1 → 77.2 (+29 eV), The Q1 quadrupole was maintained at unit resolution and Q3 at standard resolution. The Bruker MS Workstation software (Version 8.2) was used for data acquisition and processing.